AD_____

Award Number: DAMD17-02-1-0549

TITLE: Automated Analysis and Display of Temporal Sequences of Mammograms

PRINCIPAL INVESTIGATOR: Walter F. Good, Ph.D.

CONTRACTING ORGANIZATION: University of Pittsburgh
Pittsburgh, Pennsylvania 15260

REPORT DATE: July 2004

TYPE OF REPORT: Annual

PREPARED FOR: U.S. Army Medical Research and Materiel Command
Fort Detrick, Maryland 21702-5012

DISTRIBUTION STATEMENT: Approved for Public Release;
Distribution Unlimited

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 074-0188

1. AGENCY USE ONLY (Leave blank)	2. REPORT DATE	3. REPORT TYPE AND DATES COVERED
	July 2004	Annual (1 Jul 2003 - 30 Jun 2004)

4. TITLE AND SUBTITLE
Automated Analysis and Display of Temporal Sequences of Mammograms

5. FUNDING NUMBERS
DAMD17-02-1-0549

6. AUTHOR(S)

Walter F. Good, Ph.D.

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)

University of Pittsburgh
Pittsburgh, Pennsylvania 15260

E-Mail: goodwf@upmc.edu

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)
U.S. Army Medical Research and Materiel Command
Fort Detrick, Maryland 21702-5012

10. SPONSORING / MONITORING AGENCY REPORT NUMBER

11. SUPPLEMENTARY NOTES

12a. DISTRIBUTION / AVAILABILITY STATEMENT
Approved for Public Release; Distribution Unlimited

12b. DISTRIBUTION CODE

13. Abstract *(Maximum 200 Words)* *(abstract should contain no proprietary or confidential information)*

Although screening for breast cancer has been effective in detecting cancers, it is not clear that the diagnostic information present in the sequences of screening exams is currently being utilized. In an attempt to improve diagnostic accuracy, this project is adapting and integrating several novel technologies, under development in our laboratory, into a system for providing mammographers with information about changing tissue patterns, and the corresponding likelihood of malignancy, derived from temporal sequences of images. Our hypotheses are: 1) Sequences of screening mammograms contain information about tissue changes that is not otherwise being exploited in the diagnosis of breast cancer; and, 2) Changing tissue patterns can automatically be identified, and correlated with diagnostic questions. The main objectives of the project are to provide a system, which can be employed at the discretion of mammographers, to: 1) Normalize images to facilitate comparisons; 2) Apply multi-image CAD methods to identify corresponding features between images; 3) Detect and classify trends in temporal sequences; 4) Calculate and present various kinds of parameter images; and, 5) Develop a display system for efficiently presenting sequences of exams. Our expectation is that this display will improve diagnostic performance of mammographers for these cases.

14. SUBJECT TERMS	15. NUMBER OF PAGES
None Provided	40
	16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT
Unclassified	Unclassified	Unclassified	Unlimited

Table of Contents

Title: Automated analysis and Display of Temporal Sequences of Mammograms
Walter F. Good, Ph.D.

Introduction

Considerable attention has been paid to methods for improving the diagnostic accuracy and efficiency of mammographic screening procedures, including: Assessments of the importance of age of patients to be screened [1]; Appropriate screening intervals [2]; Appropriate screening techniques [3,4]; Recommended number of views per examination [5]; Single vs. double reading [6]; and, Use of prior examinations for comparison [7].

Customarily, in screening programs mammographers only compare the current exam to one previous exam (e.g., one year prior or two years prior). This is partly for efficiency but also because there is a lack of compelling evidence on how changing tissue patterns relate to the likelihood of developing a malignancy. Retrospective inspection of exams taken in years prior to the detection of a malignancy often indicates the presence of an abnormality. Such anecdotal evidence suggests that it is at least possible that series of screening exams contain additional information that could improve the accuracy of the diagnostic process.

There is also evidence that difference images in certain procedures can contain diagnostically useful information, despite the presence of substantial subtraction artifact. In the diagnosis of chest radiography, it has been demonstrated that adding temporally subtracted images to the diagnostic process can, in many situations, significantly improve diagnostic accuracy and reduce mean interpretation time [8]. Difference images obtained by temporal subtraction methods in mammography have properties somewhat similar to images obtained with bilateral subtraction images used in CAD, where it can be shown that differences contain diagnostic information [9-13]. However, like the subtraction of chest images, temporal image subtraction of mammograms is a difficult task due to tissue changes over time, inconsistencies in breast compression, skewing of the three-dimensional breast relative to the image projection during compression and lack of easy landmarks on the breast that can be used to facilitate optimal image registration. Previous attempts to develop acceptable subtraction methods for mammograms, including methods for both bilateral subtraction [9-12] and temporal subtraction [14-16], have had only partial success. As a result, the temporal subtraction of mammograms has not been employed clinically.

In an attempt to improve diagnostic accuracy of screening mammography, this project is adapting and integrating several novel technologies, under development in our laboratory, into a system for providing mammographers with information about changing tissue patterns, and the corresponding likelihood of malignancy, derived from temporal sequences of images. Specifically, these technologies include: 1) A method for locally deforming and registering mammograms [17,18]; 2) Multi-image CAD, which we have previously used to identify corresponding features in ipsilateral views [19]; and, 3) Adjustment of digitized mammograms to account nonuniform tissue thickness due to breast compression, to enhance soft display of mammograms and to enable more accurate computerized estimation of tissue composition [20]. Our hypotheses are: 1) Sequences of screening mammograms contain information about tissue changes that is not otherwise being exploited in the diagnosis of breast cancer; 2) Changing tissue patterns can automatically be identified, and correlated with diagnostic questions; and, 3) Such information can be made available to mammographers, in an efficient manner, through displays specifically designed for these image sequences.

The main objectives of this project are to provide a system, which can be employed at the discretion of mammographers, to: 1) Normalize images to facilitate comparisons; 2) Apply our previously developed multi-image CAD methods to identify corresponding features between images, and evaluate possible trends for any such features; 3) Detect and classify trends in temporal sequences (e.g., distinguish between normal and abnormal age-dependent changes, and identify those patterns of change that are associated with the development of malignancies); 4) Calculate and present various kinds of parameter images; and, 5) Develop a display system for efficiently presenting sequences of exams to mammographers, with the optional ability to emphasize trends by correcting images for differences in exposure, compression, and local misregistration. Our expectation is that this display will improve the diagnostic performance of mammographers for these cases.

Body

Progress on this project has been closely following the schedule and methods outlined in the original proposal. While many tasks have been undertaken and completed successfully, the nature of this project makes it impossible to report actual performance results at this early stage of development. The main tasks that have been completed to date are described below.

***Task 1*: Select, verify and digitize cases (Months 1 – 30)**

a. A set of 240 sequences of screening mammograms will be selected and verified

b. All verified cases will be digitized

A total of 240 cases, have been collected since this project was funded. Some of these were collected from other institutions. Of these, 125 are negative and 115 contain malignancies. All cases have been verified, digitized, and assembled into a database. The database has been anonymized by use of an honest broker system, to comply with the requirements of HIPPA. This completes the active acquisition of cases within this project.

Case selection was based on a historical prospective approach, in which we examined cases that had at least two prior screening examinations. We only considered cases for which, prior to the latest examination in which a positive abnormality may have been detected, all the mammograms for these patients had been interpreted as negative. Each case consists of two views of a single breast. Originally, we had expected to select all cases from routine mammographic examinations performed at the UPMC and Magee health care systems, but to find a sufficient quantity of positive cases having the appropriate historical sequence, we have included images from other institutions. Positive cases were drawn from cases that have undergone biopsy. Negative cases, for the most part, are cases that have not been recommended for biopsy, but in some instances, subtle negative cases, which have had biopsies that were negative for the presence of a mass or microcalcification cluster, have been selected.

Each selected film has been digitized with a 12 bit modified film digitizer (Lumisys) having an MTF of 27 and 24% at the Nyquist frequency (10 lp/mm) in the X and Y directions, respectively. The digitizer is routinely evaluated and adjusted to maintain spot size, linearity, geometric fidelity, stability and acceptable spatial noise characteristics, and to maintain a linear relationship between pixel value and optical density, as part of our quality assurance procedures.

***Task 2*: Modification of workstation for this project (Months 1 – 12)**

An existing 4-monitor display, which was developed in our laboratory, has been adapted for this study. The individual monitors (Clinton Electronics, Model DS5000) are 21 inch portrait monitors, having a resolution of 2048 x 2560 pixels, 150 ftL light output and P45 phosphor. The displays are driven by a display controller (Dome Imaging Systems, Model Md5/PCX-2) specifically designed for these monitors, and calibrated with a Barten lookup table to accommodate the contrast sensitivity of the human eye. The system has been designed to display four images of a single exam at a time, or a single image and its associated parameter images, or to display sequences of exams sequentially, at a rate determined by viewers. This task was completed in year one.

***Task 3*: Develop display software (Months 3 – 15)**

a. Display difference images

b. Display trends

The ability to deform sequentially acquired mammograms so that they can be placed in reasonable registration is a central aspect of this project. While image subtraction was addressed somewhat in year one, we continued the development of the image deformation and subtraction software in the second year. In the originally proposed method, which is described in more detail under section g of Task 4 below, we used intersections of nipple axis lines with chest walls as origins of polar coordinate systems for performing mappings between mammograms. We have demonstrated that a modification of this approach is preferable. Specifically, the origin is now defined as the midpoint of the chest wall. The line from this point, passing through the nipple serves as the base for measuring angles. But because this line is not generally perpendicular to the chest wall, a nonlinear monotonic function is used to find corresponding angles between images. Viewers, at their discretion, have the option of either subtracting consecutive images in sequence or subtracting each previous image from a specific (e.g., the most recent) image. Prior to actually performing the subtraction, images are first normalized to correct for tissue thickness, and then geometrically deformed to match a common template. The details of this procedure, and a comparison of the various techniques, have been submitted for presentation at SPIE Medical Imaging 2005.

Because the parameter images are at a much lower effective resolution that the original mammographic data, these can be displayed side-by-side. The display software has been appropriately modified to allow for alternative display formats. A cine mode has also been implemented on the display system as one means of displaying temporal sequences of image data. This can be used to display either the original mammographic data or the parameter images being generated from each mammographic image. The optimal arrangement of the displayed information will depend to a great extent on the importance of the various kinds of parameter and trend summary images being generated, and this will be determined toward the end of the project.

***Task 4*: Image analysis software development (Months 6 – 30)**

a. Correct for characteristic curve

b. Correct for breast compression

c. Skin line detection

d. Breast compression adjustment

e. Local and global composition

f. Global composition from 2 views

g. Local geometric registration

h. Histogram equalization between images
i. Modify Multi-View CAD for this project
j. Create parameter images
k. Detect and classify trends

Subtasks **a-e** and **g** were largely completed during year one.

a. Correct for characteristic curve – The characteristic curve for each particular film was obtained from the manufacturer. Although these curves are generic, they are sufficiently accurate for this application, given other imperfections in the overall imaging chain. Our routine quality assurance process on the digitizer maintains a linear relationship between pixel value and optical density. The characteristic curve correction algorithm is written to use the inverse of the digitizer characteristic curve to convert pixel values to optical density values, and then us the inverse of the film's characteristic curve to convert optical density to film exposure.

b,d. Correct for breast compression – The wide dynamic range of mammograms is caused in part by the non-uniformity of breast thickness during breast compression. To correct for this, at each pixel in a mammogram, we determine the relative thickness of the compressed breast as a function of the distance of the pixel from the skin line. Although there are computationally more efficient methods, we measure the distance of each pixel to the skin line with a simple exhaustive search. We divide the range of possible distances into a small number of intervals (typically 32 to 64 depending on image size) and for each interval, all of the tissue pixels whose distance falls within that interval are grouped. For each such group the mean and standard deviation are calculated, and an appropriate function is fitted, with constraints, to the means plus one standard deviation. This "correction" function represents the change in pixel value with respect to distance from the skin line, and actually indicates tissue thickness relative to this distance. For each pixel, the correction function is used to calculate a correction value from the pixel's distance value, and this correction value is used to normalize the pixel value. In the central regions of the breast area, the correction values are 1 and these pixels are left unchanged.

c. Skin line detection – In this project it is necessary to make geometric measurements that are referenced to the skin line and nipple. We first produce a low-resolution version of the image by averaging pixel values over a 16 16 block and then apply the Sobel gradient to the reduced-resolution image. At each point in the image, the original pixel value, the magnitude of the gradient, the direction of the gradient, the radial distance of the pixel from the center of the left edge of the film and a direction value for the surrounding neighborhood are all combined to produce a value which is proportional to the likelihood that the pixel falls on the skin line. These likelihood values are thresholded and then placed in an image where they are processed with morphological operators to eliminate isolated points. The points that remain, which are invariably on the skin line, are mapped back to the high-resolution image and used to fit a Bezier Curve to define a smooth skin line. Although failure is rather uncommon, it can happen when the tissue thickness adjustment process fails to correctly identify the skin line. This can occur for images where the skin line on the film was so dense (i.e., over an optical density of about 4.2) that the digitizer was unable to distinguish between the skin line and the image background. In this event, we do not include the case in this study.

Detection of pectoral muscle – To detect the pectoral muscle, we employ a template matching with the mammographic image, which is a standard approach, and effective in cases such as this where structures are relatively well defined. The chest wall is not normally visible in the CC view so we assume, in these cases, that it is parallel to, and 0.5 cm beyond the edge of the image.

e. Local and global composition – After correcting images for characteristic curves of film and then correcting them for nonuniformities due to breast compression, at each point in a view, we estimate % fibroglandular tissue, by deriving certain features directly from a local histogram and using a neural network, trained with values from breast MRI, to obtain a local estimate of composition. An initial global value is determined for the image by integrating over the tissue area. Tissue composition is calculated independently from CC and MLO views, and averaged to obtain the global value for the breast. The local values in each of the two views are then scaled appropriately.

f. Global composition from two views – As implemented, calculating a global composition for a breast from two views simply involves averaging the values from the CC and MLO views and then adjusting the local values in each of the individual views so that they match the averaged value. Values for individual views are found by the following procedure: 1) Correct images for characteristic curve of film; 2) Detect skin line; 3) Detect pectoral muscle; 4) Correct for breast compression; 5) Calculate fraction of fibroglandular tissue at each pixel; 6) Multiply this fraction by the breast thickness at the pixel to obtain absolute amount of fibroglandular tissue at pixel; and, 7) Integrate this over the projected area of the breast.

g. Local geometric registration – To make comparisons between sequentially acquired images more feasible, we adjust the geometry of all images being compared, so that the images can be placed in accurate registration, and we normalize grayscales so that they reflect equivalent exposure conditions. In brief, the subtraction method uses a fully automatic nonlinear transformation that can map any mammographic image onto a template or reference image while assuring concurrent and accurate registration of skin lines, nipples, pectoral muscles and nipple axis lines.

The geometric transformation is diagramed in Figure 1. The technique begins by automatically detecting pectoral muscles on MLO views (e.g., *OL* and *O'L'* in Figure 1), skin lines and nipple locations, *N* and *N'*, in Figure 1. Polar coordinate systems are established with the origins, *O* and *O'*, at the intersection of the nipple axes lines (NALs), *ON* and *O'N'*, and lines indicating the pectoral muscles, on MLO views or with lines parallel to, and *0.5 cm* beyond, the edge of the film on CC views. Tissue pixels within a mammogram are identified by a relative polar coordinate, which we define to be the angle, α, of their position vector relative to the NAL and their fractional distance between the origin and the skin line. For each pixel *P* in a reference template (e.g., left image in Figure 1), a point *P'*, having the same relative polar coordinate, is found in the image to be deformed. The pixel at location *P* in the template is then given the value of the point at *P'*, as determined by the nearest pixel or through interpolation, after adjusting for the local Jacobian of the transformation at the particular point. Clearly, this transformation maps nipples onto nipples, skin pixels onto skin pixels and chest wall pixels onto chest wall pixels, while interior pixels are deformed correspondingly.

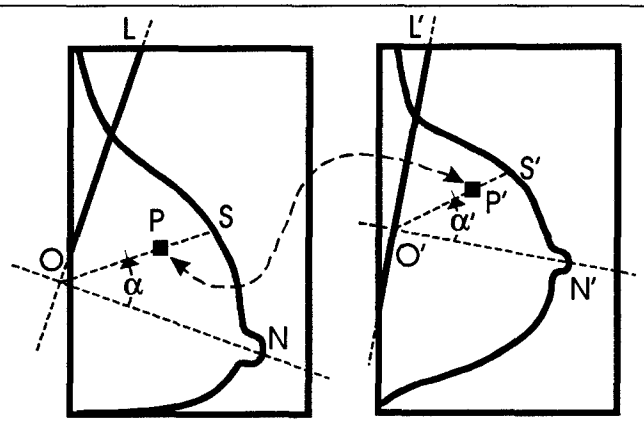


Figure 1: Relative radius polar coordinate transformation.

Grayscale adjustment – Once images have been deformed so that a reasonable correlation between pixels can be assumed, grayscales are corrected by applying a standard linear or nonlinear regression (e.g., using low degree polynomials) to pixel value pairs for the corresponding image points. Pixel values

of the deformed image are then adjusted by the regression equation, to minimize the sum-of-squares difference, with respect to admissible grayscale changes.

As stated above, the geometric registration method described above has been modified somewhat to improve performance. In the originally proposed method, we used the intersection of the nipple axis line with the chest wall as the origin of a polar coordinate system for performing a mapping between mammograms. We now have demonstrated that a modification of this approach is preferable. Specifically, the origin is now defined as the midpoint of the chest wall. The line from this point, passing through the nipple now serves as the base for measuring angles. But because this line is not generally perpendicular to the chest wall, a nonlinear monotonic function is used to find corresponding angles between images.

h. Histogram equalization between images – Once images have been deformed so that a reasonable correlation between pixels can be assumed, grayscales are corrected by fitting all pixel value pairs for corresponding image points with a 3-segment spline function consisting of two quadratic polynomials smoothly joined by a linear segment. Pixel values of the deformed image are then adjusted by the regression equation, to minimize the sum-of-squares difference, with respect to admissible grayscale changes.

i. Modify multi-view CAD for this project – Our multi-view CAD program, much of which is being adapted for use in this project, contains methods for comparing features taken at the same time from multiple views. For this project, these methods have been modified to permit their use on images taken at different times. Specifically, we included methods for applying linear regression to measure changes in features with respect to time. The main difficulty we have encountered in doing this is that our ability to find corresponding features in a sequence of images depends on our ability to deform the images so that they can be put in geometric registration.

j. Create parameter images – A general method for creating parameter images has been implemented. For temporal sequences of images, all images are deformed to the shape of the most recent image. A local feature filter is applied to each image to produce a corresponding image in which pixel values reflect the local magnitude of the particular feature in the original image. To create the trend summary image for this feature, at each pixel we calculate a correlation coefficient and the slope of the best-fit line to the sequence of parameter values. In summary images we are currently using hue to represent slope and brightness to represent correlation (with black corresponding to $r^2 = 0$).

Local features – We have defined a number of local features that could potentially be of value in identifying changing patterns in temporal sequences of mammograms. These include: local breast density, local variance, and, various measures of image texture.

As a prototype for developing these methods we used the local breast density filter that produces images somewhat similar in appearance to the original mammograms, but at a lower resolution. In summary images, locations are assigned values from a scale that ranges from cyan (falling density) to red (increasing density), where brightness indicates the degree of correlation (uncorrelated sequences of pixel values are shown in black). Other filters (e.g., local variance, local fractal dimension) have also been implemented and behave in a similar manner. We expect to continue refining these filters throughout this project.

Task 5: Initial training and testing of classifier (Months 16 – 18)

In addition to creating parameter images corresponding to individual features, we proposed to train a classifier to predict the likelihood of developing a malignancy based on actual features as well as on changes in features. Effectively this has been accomplished by adding temporal chances in features, averaged over images, to the feature vector already used in our CAD algorithms. This classification mechanism has been implemented as a neural network, and trained using the training subset of our dataset.

Because features are not independent, and the addition of temporal features has greatly increased the number of features that must be considered for input to the classifier, we will eventually use a Genetic Algorithm to define an optimal subset of features.

Task 6: Interim testing and analysis (Months 18 – 22)

We have developed methods and software to measure the contribution of each of the temporal features to the overall performance of the classifier defined above. The neural network is trained and tested on the training set of cases (half of the total dataset) using a jackknife procedure, both with and without the feature of interest. The difference in performance, with or without the feature is taken as a measure of the importance of the particular feature. Final measures of the importance of the individual temporal features will be performed once an optimal feature subset has been defined using a genetic algorithm approach.

Key Research Accomplishments

- Selected, verified and digitized 240 cases, and entered them into database
- Modified workstation for this project
- Developed display software
- Developed an innovative registration method for mammograms
- Developed Image analysis software including software to:
    - Correct for characteristic curve
    - Correct for breast compression
    - Detect Skin line
    - Adjust for breast compression
    - Calculate local and global composition
    - Global composition from two views
    - Perform local geometric registration
    - Perform histogram equalization between multiple images
- Identified local and global temporal features
- Used temporal features to create parameter images
- Derived trend summary images from the sequences of parameter images
- Developed a classification mechanism for predicting risk from temporal features

Progress on this project has been closely following the schedule and methods outlined in the original proposal. But, the nature of this project makes it impossible to report actual results at this early stage of development.

Reportable Outcomes

Manuscripts

Chang YH, Good WF, Wang XH, Glenn S. Maitz GS, Zheng B, Hardesty LA, Hakim CM, Gur D. Integrated Density of a Lesion: A Quantitative, Mammographically-derived, Invariable Measure. Med. Phys. 2003; 30(7):1805-11.

Wang XH, Chapman BE, Good WF. Evaluation of quantitative measures of breast tissue density from mammography with truth from MRI data. Proc SPIE 2003; 5032:82-89.

Good WF, Wang XH, Maitz G. Feature-based differences between mammograms. Proc SPIE 2003; 5032:919-29.

Zheng B, Leader JK, Abrams G, Shindel B, Catullo V, Good WF, Gur D. Computer-aided detection schemes: the effect of limiting the number of cued regions in each case. AJR Am J Roentgenol. 2004 Mar;182(3):579-83.

Chang YH, Good WF, Leader JK, Wang XH, Zheng B, Hardesty LA, Hakim CM, Gur D. Integrated density of a lesion: a quantitative, mammographically derived, invariable measure. Med Phys. 2003 Jul;30(7):1805-11.

Wang XH, Good WF, Fuhrman CR, Sumkin JH, Britton CA, Warfel TE, Herbert D, Gur D. Stereo Display for Chest CT. SPIE Electronic Imaging 2004; 5291: *In Press*

Wang XH, Good WF, Fuhrman CR, Sumkin JH, Britton CA, Warfel TE, Gur D. Projection Models for Stereo Display of Chest CT. SPIE Medical Imaging 2004; 5367: *In Press*

Wang XH, Good WF, Fuhrman CR, Sumkin JH, Britton CA, Warfel TE, Gur D. Stereo Display for Chest CT. RSNA Dec 2003;

Databases

We have established a database of verified cases, where each case consists of a sequence of screening mammography exams. While we will continue to accrue cases to this database, it is currently in a form that can be used for analysis.

Proposals applied for

US Army MRMC 2003	Assessing the Interdependence of Tumor Risk, Conspicuity and Breast Tissue Characteristics Derived from Mammography
US Army MRMC 2003	Computerized Analysis and Display of Contralateral Breast Asymmetry
US Army MRMC 2003	Improving Breast Cancer Assessment Through More Accurate Measurement of Mass Size and Growth Rate
Komen 2002	Integrated Density – A Quantitative Measure of Breast Lesion
NIH 2004	Improving Early Detection of Breast Cancer

Conclusions

Progress on this project has been closely following the schedule and methods outlined in the original proposal. To this point, all of our effort has been aimed at assembling a database of images, to be used in subsequent years for analysis, and in developing the display and software algorithms. While many such tasks have been undertaken and completed successfully, the nature of this project makes it impossible to report actual performance results until the final year of the project.

The workstation we have developed will make it possible for radiologists to efficiently view a woman's entire mammographic history, as well as view summarized presentations of that history. We expect that the viewing of this additional data will increase radiologists' understanding of the kinds of changes that are occurring in breasts, and the rapidity with which these changes are occurring. Ultimate, this will likely result in increased efficacy for screening mammography.

During the current year of the project, we will continue addressing the remaining tasks. These include:

Task 4: Image analysis software development (Months 6 – 30)
k. Detect and classify trends
Task 7: Final training and testing of classifier (Months 24-30)
Task 8: Final processing of all cases (Months 30 – 34)
Task 9: Data analysis and final report (Months 30-36)

References

1. Feig SA, D'Orsi CJ, Hendrick RE. American college of radiology guidelines for breast cancer screening. *AJR* **1998;** 171:29-33.

2. Gordon R. Detection of breast cancer at a small size can reduce the likelihood of metastatic spread: quantitative analysis. *Acad Radiol* **1997;** 4:8-12.

3. Hanchak NA, Kessler HB, MacPerson S. Screening mammography: experience in a health maintenance organization. *Radiology* **1997;** 205:441-445.

4. Moskowitz M. Retrospective reviews of breast screening: what do we really learn from them? *Radiology* **1996;** 199:615-620.

5. Sickles EA. Findings at mammographic screening on only one standard projection: outcomes analysis. *Radiology* **1998;** 209:471-475.

6. Thurfjell EL, Lernevall KA, Taube AS. Benefit of independent double reading in a population-based mammography screening program. *Radiology* **1994;** 191:241-244.

7. Callaway MP, Boggis CR, Astley SA. Influence of previous films on screening mammographic interpretation and detection of breast carcinoma. *Clin Radiol* **1997;** 52:527-529.

8. Difazio MC, MacMahon H, Xu XW, Tsai P, Shiraishi J, Armato SG, Doi K. Digital chest radiography: effect of temporal subtraction images on detection accuracy. *Radiology* **1997**; 202:447-452.

9. FF Yin, ML Giger, K Doi, CE Metz, CJ Vyborny, RA Schmidt. Computerized detection of masses in digital mammograms: analysis of bilateral subtraction images. *Med Phys* **1991**; 18:955-963.

10. FF Yin, ML Giger, CJ Vyborny, K Doi, RA Schmidt. Comparison of bilateral-subtraction and single-image processing techniques in the computerized detection of mammographic masses. *Invest Radiol* **1993**; 6:473-481.

11. FF Yin, ML Giger, K Doi, CJ Vyborny, RA Schmidt. Computerized detection of masses in digital mammograms: Automated alignment of breast images and its effect on bilateral-subtraction technique. *Med Phys* **1994**; 21:445-452.

12. B Zheng, YH Chang, D Gur. Computerized detection of masses from digitized mammograms: Comparison of single-image segmentation and bilateral image subtraction. Acad Radiol **1995**; 2:1056-106

13. Mendez AJ, Tahocas PG, Loda MJ. Computer-aided diagnosis: automatic detection of malignant masses in digital mammograms. *Med Phys* **1998;** 25:957-964.

14. M Sallam, K Bowyer. Registering time sequences of mammograms using a two-dimensional image unwarping technique. SIWDM **1994.**

15. D Brzakovic, N Vujovic, M Neskovic, P Brzakovic, K Fogarty. Mammogram analysis by comparison with previous screenings. SIWDM **1994.**

16. S. Sanjay-Gopal, HP Chan, N Petrick, TE Wilson, B Sahiner, MA Helvie, MM Goodsitt. A regional mammogram registration technique for automated analysis of interval changes of breast lesions. Proc. SPIE 3338:118-123.

17. Good WF, Zheng B, Chang YH, Wang XH, Maitz G. Generalized procrustean image deformation for subtraction of mammograms. *Proc SPIE on Medical Imaging* **1999;** 3666-167.

18. Good WF, Zheng B, Chang YH, Wang XH, Maitz G. Image modification for displaying temporal sequences of mammograms. *Proc SPIE on Medical Imaging* **2000;** 3976 (in press).

19. Good WF, Zheng B, Chang YH, Wang XH, Maitz GS, Gur D. Multi-image CAD employing features derived from ipsilateral mammographic views. *Proc SPIE on Medical Imaging* **1999;** 3661-47.

20. Wang XH, Zheng B, Chang H, Good WF. Correction of digitized mammograms to enhance soft display amd tissue composition measurement. *Proc SPIE on Medical Imaging* **2001**; Vol 4319.

Appendices

1. Chang YH, Good WF, Wang XH, Glenn S. Maitz GS, Zheng B, Hardesty LA, Hakim CM, Gur D. Integrated Density of a Lesion: A Quantitative, Mammographically-derived, Invariable Measure. Med. Phys. 2003; 30(7):1805-11.

2. Wang XH, Chapman BE, Good WF. Evaluation of quantitative measures of breast tissue density from mammography with truth from MRI data. Proc SPIE 2003; 5032:82-89.

3. Good WF, Wang XH, Maitz G. Feature-based differences between mammograms. Proc SPIE 2003; 5032:919-29.

Integrated density of a lesion: A quantitative, mammographically derived, invariable measure

Yuan-Hsiang Chang,[a] Walter F. Good, Joseph K. Leader, Xiao-Hui Wang, Bin Zheng, Lara A. Hardesty, Christiane M. Hakim, and David Gur
From the Department of Radiology, University of Pittsburgh and Magee-Womens Hospital, Pittsburgh, Pennsylvania 15213

(Received 8 January 2003; revised 1 April 2003; accepted for publication 28 April 2003; published 25 June 2003)

A method for quantitatively estimating lesion "size" from mammographic images was developed and evaluated. The main idea behind the measure, termed "integrated density" (ID), is that the total x-ray attenuation attributable to an object is theoretically invariant with respect to the projected view and object deformation. Because it is possible to estimate x-ray attenuation of a lesion from relative film densities, after appropriate corrections for background, the invariant property of the measure is expected to result in an objective method for evaluating the "sizes" of breast lesions. ID was calculated as the integral of the estimated image density attributable to a lesion, relative to surrounding background, over the area of the lesion and after corrections for the nonlinearity of the film characteristic curve. This effectively provides a measure proportional to lesion volume. We computed ID and more traditional measures of size (such as "mass diameter" and "effective size") for 100 pairs of ipsilateral mammographic views, each containing a lesion that was relatively visible in both views. The correlation between values calculated for each measure from corresponding pairs of ipsilateral views were computed and compared. All three size-related measures (mass diameter, effective size, and ID) exhibited reasonable linear relationship between paired views ($r^2>0.7, P<0.001$). Specifically, the ID measures for the 100 masses were found to be highly correlated ($r^2=0.9, P<0.001$) between ipsilateral views of the same mass. The correlation increased substantially ($r^2=0.95$), when a measure with linear dimensions of length was defined as the cube root of ID. There is a high degree of correlation between ID-based measures obtained from different views of the same mass. ID-based measures showed a higher degree of invariance than mass diameter or effective size. © *2003 American Association of Physicists in Medicine.*
[DOI: 10.1118/1.1582571]

Key words: breast, imaging, masses, mammography, quantitative analysis

I. INTRODUCTION

Mammography has been frequently recommended as a routine screening tool to detect breast cancers at an earlier stage.[1] While it has been shown that mammographic screening can substantially reduce cancer-related mortality and morbidity,[2,3] identifying breast cancers in the screening environment with high sensitivity and specificity is a difficult task due to the low expected cancers detected in a large volume of cases and the complex patterns as depicted on mammograms.[4,5] Detection and diagnostic accuracy of breast cancers using mammograms can be improved using quantitative analysis of masses.[6–9] Studies of cases with prior examinations demonstrated that a change in the density or contour of a mass over time can be an indicative sign of malignancy.[10] Other studies showed that the change in mass size was one of the dominant factors in determining breast cancer prognosis.[11,12] However, inter- and intra-observer variability, when visually and subjectively describing a mass or its change over time between consecutive examinations, makes this assessment quite difficult and often unreliable.[13,14]

Quantitative measurements and analyses of masses have been widely used in computer-aided detection and characterization (discrimination) schemes.[15,16] A large number of features (including morphologic, texture, and density based[17,18]) have been investigated in an attempt to quantitatively and objectively represent masses. Currently, there is no standard that defines a mass and its surrounding background. Without such a standard, studies have shown that measured contrast values of a mass region could change substantially if the definition of the surrounding background of a mass was varied.[19] Therefore, measurements of mass contrast and other related features are frequently scheme dependent. In addition, due to the wide variation of tissue presentation resulting from breast compression, image projection, and field nonuniformity, a large number of computed image-based features of a mass, as measured from different images, is not a constant.[20,21] Therefore, it would be desirable to define a measure of a mass that is invariant to tissue deformation and the projection view. It has been theoretically shown that under a few conditions, the total attenuation (termed here integrated density or ID) of an object is an invariant quantity with respect to geometrical deformation of an object in two-dimensional projected images (such as mammograms).[22] There is no experimental validation that in fact ID is an invariant measure.[22] In this study, we experimentally evalu-

ate integrated density (ID) as an invariant measure representing masses depicted on mammograms. After correcting for film nonlinearity, and estimating the density of underlying tissue, we assess whether ID as computed for the same 100 masses depicted during the same examination on both CC and MLO views is invariant. The purpose of this preliminary study is intended only to assess the degree of ID invariance with respect to breast compression and projection views, and not changes over time.

II. MATERIALS AND METHODS

A. Defining integrated density

Assuming a monoenergetic x-ray source, for a mass present in a breast, the quantity

$$\int\int\int_V \mu_M \, dV,$$

where V is the volume occupied by the mass and μ_M is its local x-ray attenuation coefficient, is invariant, but it cannot be measured directly from the image. However, if it is assumed that the attenuation coefficient of the breast tissue surrounding the mass (μ_N) can be estimated with reasonable accuracy, then the quantity

$$\int\int\int_V (\mu_M - \mu_N) dV = \int\int\int_V \Delta\mu \, dV,$$

remains essentially invariant and can be approximated from an image.

To demonstrate this, consider the density of a single pixel within the image projection of the mass. In the case of a mammographic film image, for which the digitized density values have been adjusted for the film's characteristic curve, the corrected density at an image pixel, $D\,(u,v)$, can be calculated as

$$\begin{aligned} D(u,v) &= \text{Const} + \gamma \log\left[E \cdot \exp\left(-\int_L \mu_N dx - \int_{L_M} (\mu_M - \mu_N) dx \right) \right] \\ &= \text{Const} + \gamma \log E - \gamma \int_L \mu_N dx - \gamma \int_{L_M} (\mu_M - \mu_N) dx, \end{aligned}$$

where E is the overall exposure, γ is the film's gamma, the first integral is taken over the whole breast tissue (L) as projected onto the image, and the second integral is taken over the mass region (L_M) as projected onto the image. The first two terms on the left of the expanded expression correspond to background density, D_{BKG}, of the film (i.e., film base plus nonattenuated exposure including scatter contribution). The third term is the reduction in density resulting from surrounding tissues, D_{NORMAL}, and the final term is the reduction in density attributed to the presence of the mass. This pixel density can be rewritten as

$$D(u,v) = D_{\text{BKG}} - D_{\text{NORMAL}} - \gamma \int_{L_M} \Delta\mu \, dx,$$

or, rearranging this equation,

$$\gamma \int_{L_M} \Delta\mu \, dx = D_{\text{BKG}} - D_{\text{NORMAL}} - D(u,v).$$

If we integrate this over the region, R, defined by the projection of the mass, we obtain

$$\begin{aligned} &\int\int_R \left(\gamma \int_{L_M} \Delta\mu \, dx \right) du \, dv \\ &= \int\int_R (D_{\text{BKG}} - D_{\text{NORMAL}}) du \, dv - \int\int_R D(u,v) du \, dv, \end{aligned}$$

and simplifying the left-hand side,

$$\begin{aligned} \gamma \int\int\int_V \Delta\mu \, dV &= \int\int_R (D_{\text{BKG}} - D_{\text{NORMAL}}) du \, dv \\ &\quad - \int\int_R D(u,v) du \, dv. \end{aligned}$$

The left-hand side of this expression is the film's γ times a quantity that is expected to be essentially invariant; hence, it should be invariant for a particular type of film (in the "linear" range). Thus, we define ID as

$$\text{ID} = \gamma \int\int\int_V \Delta\mu \, dV,$$

which can be approximated as

$$\text{ID} \approx (D_{\text{BKG}} - D_{\text{NORMAL}} - D_{\text{MASS}}) A_R = C \cdot A_R.$$

Therefore, ID is represented by the area of the mass, A_R, multiplied by the average contrast difference between the mass and the surrounding tissue, where the mass area is defined to be the projected area associated with the interior of the mass and the average mass contrast, C, is defined as the average difference in linearized densities between the underlying tissue background and the mass itself.[19] It should be emphasized that x-ray beam hardening is ignored in this simplification and scatter radiation is assumed to contribute a relatively uniform distribution in both the mass and background areas. Hence, it can be represented as a constant in the background term. Determining the relative change in log-exposure, from film density in mammograms, involves approximating the density due to the combination of background (including scatter) and normal tissues (i.e., $D_{\text{BKG}} - D_{\text{NORMAL}}$), which would be present if the lesion did not exist. This can be estimated as the mean pixel value of the area outside the lesion. It should be noted that such estima-

tion is based on the assumption that the "normal" tissue is distributed "uniformly" over the area of a projected lesion and the surrounding area. Therefore, measurement of the surrounding area would closely resemble the measurement of the underlying tissue if the lesion was absent. Because of the assumptions mentioned above, it was the intent of this study to assess the extent to which ID as simply computed from the image remains invariant to deformation or projection view.

Film characteristic curve linearization: Digitized film density values were corrected for the characteristic curve to produce values that were linearly proportional to log-exposure (in the linear range). First, the laser film digitizer was routinely calibrated to assure that film optical density (OD) was linearly translated into digitized pixel values in the density range of interest. Second, film OD values were corrected so that they were linearly proportional to log-exposure units. A generic curve was used for this purpose using the data for the specific mammography film used in our facility (Eastman Kodak Min-R 2000 film).[23] For computational ease, the generic curve was represented by a spline function.[23] To define ID in terms of more familiar "density" units, rather than log-exposure units, we converted back exposure values to linearized density values by fitting a straight line to this curve. As a result, for each OD value, we computed the corresponding log-exposure unit using the spline function and then converted it to a "linearized" OD value using the "fitted" line.

Delineating mass regions: For each mass, the corresponding pair of mammographic views was reviewed by experienced mammographers who initially identified a central point in the projection of the mass as depicted in two projections. A semi-automatic delineation of mass boundaries was then performed using a region growing routine similar to the technique described by Matsumoto *et al.*[24] For each projection, the location representing the central point (pixel) of the mass was first identified on digitized mammograms. Based on the initial location, the algorithm automatically determined a "transition layer," where a substantial change in region growth and margin irregularity was observed. All pixels within the identified region boundary were considered to be in the region (R) depicting the mass. Mass delineation could affect the results since both the area of the mass (A_R) and the average density within the mass (D_{MASS}) are used to compute ID.

Measuring mass background: Once a mass region (R) had been delineated, the geometric center-of-mass of the lesion was computed, and the maximum distance of this point to the lesion boundary was determined. For background density estimation, a region was defined as all tissue regions outside the lesion, within a circle centered at the center of mass (R'), with a radius difference (Δ radius) longer than the maximum distance to the lesion's boundary. Linearized density values within the background region (R') were averaged.

To estimate the density of normal tissue in the areas where the mass overlaps in the projection image, we used the average density value of the surrounding area outside the mass. The radius difference was initially chosen as Δ radius $= 1$ cm for the background estimation. However, since the definition of the background region can affect the computed ID, we investigated this issue in the following manner. Based on the distribution of mass diameters in our database, we selected four values for the radius differences (Δ radius $= 5$ mm, 7.5 mm, 10 mm, or 12.5 mm) to be considered as the background area. IDs were estimated from the CC and MLO views using the different background definitions, and the correlations between computed ID values for the corresponding views were computed.

FIG. 1. An example of a pair of images depicting a mass in two views. The mass is relatively visible in both the CC (left) and MLO (right) views. Superimposed on each view are the mass center ("+"), the computed mass boundary, and the corresponding mass background (circular region excluding the interior of the mass region).

ID computation for each mass: The area of the mass, A_R, was obtained by counting the number of pixels inside the region R and converting it to an area (one pixel represented an area of $100 \times 100\ \mu\mathrm{m}^2$ or $0.01\ \mathrm{mm}^2$). Average mass contrast, C, was computed as the difference in average density values within the regions R and R'. Finally, ID is the product of the mass area and the average mass contrast (i.e., $\mathrm{ID} \approx C \cdot A_R$). This process was performed independently for each mammographic view.

Figure 1 demonstrates an example of the regions analyzed in one case. The mass was clearly visible in both the CC view and MLO views. In each view, an irregular boundary marks the mass region as automatically determined by the software. The background region of surrounding tissue that was used in the analysis is also shown.

B. Other relevant measures

In addition to the computed ID, two mammographically based measures were derived for each projection of a mass, as follows: (1) Mass diameter was defined as the maximum diameter (or longest axis) of the mass as depicted in the image (in mm); and (2) effective size was defined as the squared root of the product of maximum and minimum di-

ameters (longest axis and shortest axis) and was also measured in mm.[25] These measures were used for comparison with ID, as these are frequently used in the clinical environment and in CAD assessments.

C. Dataset

A total of 100 verified cases were selected from our database of patients who have undergone screening mammography in one of our clinical facilities. Each case included an ipsilateral pair of craniocaudal (CC) and mediolateral oblique (MLO) mammographic images. CC and MLO views of the same breast from the same examination were used. Cases were selected only if a well-defined mass was depicted on both views. Of the 100 selected masses, 64 were malignant and 36 benign. All films were digitized using a laser film digitizer (Lumisys, Eastman Kodak Co., Rochester, NY) at $100\times100\ \mu m^2$ pixel size and 12-bit contrast resolution. The laser film digitizer was routinely calibrated to assure that film optical density (OD) is linearly translated into digital pixel values in the range of 0.2 to 3.8 OD (1 pixel value $=0.001$ OD). The three measures (ID, diameter, and effective size) were estimated for both views of each of the 100 masses in the dataset.

D. Evaluation

In this study we compared the results of the three measures as computed independently for each of the two views. For each measure, we computed the correlation (Pearson's r^2) between values computed from the CC views and those from the MLO views. The results for each measure were also plotted in corresponding scattergrams.

Because of the units associated with these measures, it is suboptimal to compare directly the correlations for ID with those computed for the other two measures. ID is proportional to volume, while the other two measures are defined in units of lengths. Therefore, we defined a cube root of ID as a more appropriate measure for comparisons (which is given in unit of length) and report the results of this measure, as well.

To evaluate size-dependent differences between paired measurements obtained from the two views for each of the 100 masses, we divided the database into three subsets (<10 mm, 10–20 mm, and >20 mm). The absolute values of the differences and the range were evaluated for each of the subsets.

Geometric eccentricity of a mass is one factor that could affect the results. Therefore, we categorized masses into two groups by their eccentricity, and repeated the analyses for each of the subgroups. We classified cases into the subgroups using the two diameters d_{CC} and d_{MLO}, obtained from the two views, and computed a ratio $\hat{e}$ for each mass as: $\hat{e}=\max(d_{CC}/d_{MLO}, d_{MLO}/d_{CC})$. All masses for which $\hat{e}\leq1.1$, were assigned to one group (more "spherical") and the remaining masses (more "eccentric") were assigned to the other group. Masses with substantially different diameters as depicted on the CC and MLO views typically exhibit high eccentricity. Thus, a relatively low correlation for diameter or effective size was expected in this group.

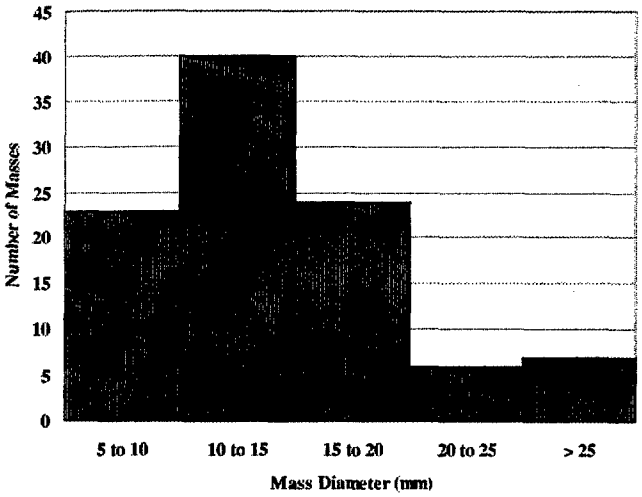


FIG. 2. Distribution of the mass sizes for the 100 masses used in this study. Mass size was determined by averaging maximum diameters (mm) obtained from the two ipsilateral views (CC and MLO).

III. RESULTS

Figure 2 shows the distribution of mass sizes. The size of each mass was determined by averaging the two maximum diameters obtained from the two views (CC and MLO). As can be seen from the figure, the 100 selected masses ranged in sizes from relatively small (<10 mm) to quite large (>25 mm).

Table I shows the linear correlation coefficients (Pearson's r^2) between matched pairs of measures (i.e., mass diameter, effective size, mass area, mass contrast, and ID) from the two views. Figure 3 shows the corresponding scattergrams for (a) mass diameter, (b) effective size, (c) ID, and (d) cube root of ID. The size-related measures (mass diameter and effective size) were found to follow a reasonable linear relationship ($P<0.001$). Despite the relatively weaker correlation exhibited by the mass contrast ($r^2=0.45$), the ID measure, which is the product of mass area and relative contrast, highly correlated between paired measurements ($r^2=0.9, P<0.001$). Figure 3 also demonstrates the scattergram of the cube root of ID ($\sqrt[3]{mm^2\cdot\Delta OD}$), which allows for a more valid comparison with the mass diameter and effective size

TABLE I. Linear correlation coefficients (*Pearson's* r^2) of various mass size-related measures (mass diameter, effective size, mass area, mass contrast, and ID) as measured from the CC and MLO view using 100 masses for the assessment.

	Mass diameter	Effective size	Mass area	Mass contrast	ID
Correlation coefficient r^2	0.74	0.79	0.83	0.45	0.90

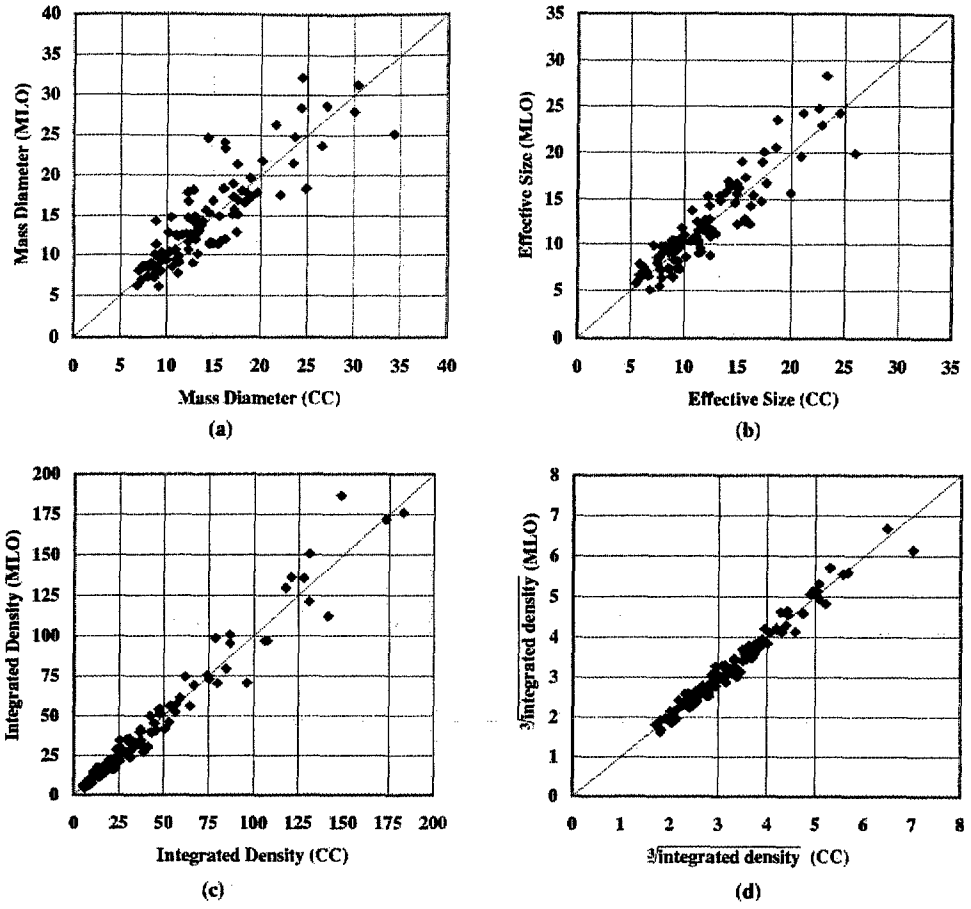


FIG. 3. Scattergrams of the quantitative measures (a) mass diameter, (b) effective size, (c) ID, and (d) cube root of ID, as computed from paired ipsilateral views of 100 masses. The diagonal line represents identical measures.

since it is represented by a similar dimension. The correlation coefficient for the data shown for this measure is substantially higher $r^2=0.95$.

Table II shows the ranges of the absolute differences between paired measures obtained from the two views for the different subsets of mass sizes. Three subsets (mass diameter<10 mm, 10–20 mm, and >20 mm) are shown. Differences for all measures increase as the mass diameter increase. Measured eccentricity values were 1.14 ± 0.12, 1.17 ± 0.16, and 1.19 ± 0.15 for the three groups, respectively.

Table III shows the linear correlation coefficients of the measures, for each of the subgroups of masses based on eccentricity. As can be seen from this table, both ID and cube root of ID correlated well for the different types of masses, those with low or high eccentricity. The other two measures

TABLE II. Absolute values of differences between paired measures obtained from two corresponding views for the three subsets of masses segmented by mass size.

	Mass diameter		
	<10 mm	10–20 mm	>20 mm
Number of masses	23	64	13
Difference in mass diameter Δmm	1.1 ± 0.8	2.2 ± 1.9	4.0 ± 2.9
Difference in effective size Δmm	1.1 ± 0.7	1.6 ± 1.0	2.6 ± 1.8
Difference in ID $\Delta(mm^2\cdot\Delta OD)$	2.2 ± 1.5	5.1 ± 5.7	20.2 ± 26.9
Difference in cube root of ID $\Delta(\sqrt[3]{mm^2\cdot\Delta OD})$	0.13 ± 0.1	0.14 ± 0.07	0.21 ± 0.22

TABLE III. Linear correlation coefficients (*Pearson's* r^2) between corresponding paired views of the same masses for the two groups of cases segmented by eccentricity.

	Eccentricity ($\hat{e}$)	
	≤1.1	>1.1
Number of masses	44	56
Mass diameter (r^2)	0.98	0.58
Effective size (r^2)	0.85	0.74
ID (r^2)	0.92	0.90
Cube root of ID (r^2)	0.95	0.95

TABLE IV. Linear correlation coefficients (*Pearson's* r^2) between corresponding paired views for the malignant and benign masses.

	Malignant masses	Benign masses
Number of masses	64	36
ID (r^2)	0.90	0.92

exhibited lower correlations for the subset of masses with high eccentricity ($\hat{e}$).

Table IV shows the correlation of ID values between CC and MLO views for malignant (64) and benign (36) cases, respectively. While malignant masses generally exhibit more irregular boundary, our experimental results showed in this group of well-defined masses that the ID correlations were similar (0.90 and 0.92 for malignant and benign masses, respectively).

Table V shows the range of averaged ID values and correlation coefficients between paired CC and MLO views when different areas were used for background definition (i.e., Δ radius = 5 mm, 7.5 mm, 10 mm, or 12.5 mm, respectively). Although the ranges of ID measurements varied with respect to background definitions, ID values were well correlated between the paired views in all measurements.

IV. DISCUSSION

We evaluated a quantitative measure for estimating the "sizes" of masses as depicted in mammograms. For 100 verified masses, ID was "estimated" as the product of its area and average linearized contrast. ID was found to be relatively "invariant" between the two views (CC and MLO). It was shown to be a better measure than the others tested in this study, and in particular it is better than the other measures for the subset of masses that are more eccentric. When ID was transformed to a measure with a unit comparable to length (by taking the cube root), its performance increased substantially, resulting in a correlation coefficient of $r^2 = 0.95$.

Computed ID values are dependent on the segmentation method and the definition of the surrounding background region used. Therefore, ID remains scheme dependent. However, despite a reasonably wide distribution of mass sizes and shapes and the background areas evaluated, our results suggest that ID is reasonably invariant with respect to the image projection (view).

TABLE V. Range of averaged measured ID values and the correlation coefficients between paired views for different background definitions. Δ *radius* = 5 *mm*, 7.5 mm, 10 mm, and 12.5 mm were used.

	Radius difference (Δ *radius*) for background definition			
	5 mm	7.5 mm	10 mm	12.5 mm
Range of averaged IDs ($mm^2 \cdot \Delta OD$)	40.7±43.7	44.5±47.6	48±51.6	51.2±55.7
Correlation (*Pearson's* r^2)	0.92	0.91	0.90	0.88

Because of the presence of the mass, one can only approximate the measurement from the surrounding area using the assumption that the normal tissue in the area of the projected mass and surrounding area are the same. This assumption is but one source of error in the computed ID value that could affect its invariance properties. This preliminary study was intended mainly to assess the performance of ID as an invariant measure of mass size. Hence, we included only cases with relatively well-defined masses that were clearly visible on both views. As a result, the reported correlation coefficients for all measures applied to our dataset are likely to overstate the performance of such a measure in a dataset that includes more subtle or somewhat ill-defined cases.

The assumptions that the measure will not be significantly affected by scatter, film gamma, and exposure factors (e.g., geometry, kVp, filtration, and field uniformity[21]) seem to be reasonable for this purpose, but the desired characteristic (invariance) will have to be experimentally confirmed for different experimental conditions.

Our findings suggest that it may be possible to achieve a relatively reproducible measure for a given mass over a rather wide range of conditions and different mammographic views. We appreciate the fact that the definition of some of the measurements of interest in this work were simplified and could be more accurately described. However, the intent was to develop a relatively easy measure to compute, perhaps at the cost of being somewhat less rigorous and precise.

V. CONCLUSION

We have developed a method for deriving a quantitative measure of lesion "size," termed integrated density or ID, that was found to be reasonably invariant between paired projection views of the same breast. Our experimental results in a set of 100 well-defined masses demonstrated a high degree of correlation between ID-based measures obtained from ipsilateral paired views of the same breast. ID-based measures showed a higher degree of correlation when compared with other traditional size-related measures, such as mass diameter or effective size.

ACKNOWLEDGMENTS

This work is supported in part by Grants Nos. CA85241, CA77850, and CA80836 from the National Cancer Institute, National Institutes of Health, Grant No. IMG-2000-362 from the Susan G. Komen Breast Cancer Foundation, and also by the US Army Medical Research Acquisition Center, 820 Chandler Street, Fort Detrick, MD 21702-5014 under Contract No. DAMD17-02-1-0549. The content of the information contained here does not necessarily reflect the position or the policy of the government, and no official endorsement should be inferred.

[a]Electronic mail: ychang@mail.magee.edu

[1] S. A. Feig, C. J. D'Orsi, R. E. Hendrick, V. P. Jackson, D. B. Kopans, B. Monsees, E. A. Sickles, C. B. Stelling, M. Zinninger, and P. Wilcox-Burchalla, "American College of Radiology guidelines for breast cancer screening," AJR, Am. J. Roentgenol. **171**, 29–33 (1998).

[2] S. A. Feig, "Increased benefit from shorter screening mammography intervals for women ages 40–49 years," Cancer **80**, 2035–2039 (1997).
[3] L. Tabar, B. Vitak, H. H. Chen, M. F. Yen, S. W. Duffy, and R. A. Smith, "Beyond randomized trials: organized mammographic screening substantially reduces breast cancer mortality," Cancer **91**, 1724–1731 (2001).
[4] S. K. Goergen, J. Evans, G. P. Cohen, and J. H. MacMillan, "Characteristics of breast carcinomas missed by screening radiologists," Radiology **204**, 131–135 (1997).
[5] R. L. Birdwell, D. M. Ikeda, K. F. O'Shaughnessy, and E. A. Sickles, "Mammographic characteristics of 115 missed cancers later detected with screening mammography and the potential utility of computer-aided detection," Radiology **219**, 192–202 (2001).
[6] J. E. Meyer and D. B. Kopans, "Stability of a mammographic mass: a false sense of security," AJR, Am. J. Roentgenol. **137**, 595–598 (1981).
[7] E. A. Sickles, "Breast masses: mammographic evaluation," Radiology **173**, 297–303 (1989).
[8] H. K. Hussain, Y. Y. Ng, C. A. Wells, I. B. Nockler, O. M. Curling, R. Carpenter, and N. M. Perry, "The significance of new densities and microcalcification in the second round of breast screening," Clin. Radiol. **54**, 243–247 (1999).
[9] S. Sanjay-Gopal, H. P. Chan, T. Wilson, M. Helvie, N. Petrick, and B. Sahiner, "A regional registration technique for automated interval change analysis of breast lesions on mammograms," Med. Phys. **26**, 2669–2679 (1999).
[10] G. Hermann, R. J. Keller, P. Tartter, I. Bleiweiss, and J. G. Rabinowitz, "Interval changes in nonpalpable breast lesions as an indication of malignancy," Can. Assoc. Radiol. J. **46**, 105–110 (1995).
[11] C. L. Carter, C. Allen, and D. E. Henson, "Relation of lesion size, lymph node status, and survival in 24,740 breast cancer cases," Cancer **63**, 181–187 (1989).
[12] G. D'Eredita', C. Giardina, M. Martellotta, T. Natale, and F. Ferrarese, "Prognostic factors in breast cancer: the predictive value of the Nottingham Prognostic Index in patients with a long-term follow-up that were treated in a single institution," Eur. J. Cancer **37**, 591–596 (2001).
[13] J. A. Baker, P. J. Kornguth, and C. E. Floyd, Jr., "Breast imaging reporting and data system standardized mammography lexicon: observer variability in lesion description," Am. J. Roentgenol. **166**, 773–778 (1996).
[14] L. Liberman and J. H. Menell, "Breast imaging reporting and data system (BI-RADS)," Radiol. Clin. North Am. **40**, 409–430 (2002).
[15] C. J. Vyborny, M. L. Giger, and R. M. Nishikawa, "Computer-aided detection and diagnosis of breast cancer," Radiol. Clin. North Am. **38**, 725–740 (2000).
[16] I. Leichter, S. Buchbinder, P. Bamberger, B. Novak, S. Fields, and R. Lederman, "Quantitative characterization of mass lesions on digitized mammograms for computer-assisted diagnosis," Invest. Radiol. **35**, 366–372 (2000).
[17] B. Sahiner, H. P. Chan, N. Petrick, M. A. Helvie, and M. M. Goodsitt, "Computerized characterization of masses on mammograms: The rubber band straightening transform and texture analysis," Med. Phys. **25**, 516–526 (1998).
[18] L. Li, Y. Zheng, L. Zhang, and R. A. Clark, "False-positive reduction in CAD mass detection using a competitive classification strategy," Med. Phys. **28**, 250–258 (2001).
[19] B. Zheng, Y. H. Chang, and D. Gur, "On the reporting of mass contrast in CAD research," Med. Phys. **23**, 2007–2009 (1996).
[20] W. F. Good, B. Zheng, Y. H. Chang, X. H. Wang, and G. Maitz, "Procrustean image deformation for bilateral subtraction of mammograms," Proc. SPIE **3661**, 1562–1573 (1999).
[21] O. Pawluczyk and M. J. Yaffe, "Field nonuniformity correction for quantitative analysis of digitized mammograms," Med. Phys. **28**, 438–444 (2001).
[22] J. M. Fitzpatrick, "The existence of geometrical density—Image transformations corresponding to object motion," Comput. Vis. Graph. Image Process. **44**, 155–174 (1988).
[23] B. J. McParland, "A comparison of two mammography film-screen combinations designed for standard-cycle processing," Br. J. Radiol. **72**, 73–75 (1999).
[24] T. Matsumoto, H. Yoshimura, K. Doi, M. L. Giger, A. Kano, H. MacMahon, K. Abe, and S. M. Montner, "Image feature analysis of false-positive diagnoses produced by automated detection of lung nodules," Invest. Radiol. **27**, 587–597 (1992).
[25] R. M. Nishikawa, M. L. Giger, K. Doi, C. E. Metz, F. F. Yin, C. J. Vyborny, and R. A. Schmidt, "Effect of case selection on the performance of computer-aided detection schemes," Med. Phys. **21**, 265–269 (1994).

Evaluation of quantitative measures of breast tissue density from mammography with truth from MRI data

Xiao Hui Wang, Brian E. Chapman, Cynthia A. Britton, Saraswathi K. Golla, Louisa P. Wallace, Walter F. Good.
Department of Radiology, University of Pittsburgh, Pittsburgh, PA 15261

ABSTRACT

Breast tissue density is one of the most cited risk factors in breast cancer development. Nevertheless, estimates of the magnitude of breast cancer risk associated with density vary substantially because of the inadequacy of methods used in tissue density assessment (e.g., subjective and/or qualitative assessment) and lack of a reliable gold standard. We have developed automated algorithms for quantitatively measuring breast composition from digitized mammograms. The results were compared to objective truth as determined by quantitative measures from breast MR images, as well as to subjective truth as determined by radiologists' readings from digitized mammograms using BI-RAD standards. Higher linear correlation between estimates calculated from mammograms using the methods developed herein and estimates derived from breast MR images demonstrates that the mammography-based methods will likely improve our ability to accurately determine the breast cancer risk associated with breast density. By using volumetric measures from breast MR images as a gold standard, we are able to estimate the adequacy and accuracy of our algorithms. The results can be used for providing a calibrated method for estimating breast composition from mammograms.

Keywords: breast tissue composition, mammography, MRI breast imaging

1. INTRODUCTION

Assessing risk factors is one crucial step involved in breast cancer detection and early prevention. Because of its apparent association with breast cancer development, and possible masking effects of early stage cancerous changes, dense breast tissue is often cited as one breast cancer risk factor, although the magnitude of the risk is still debatable [1-9].

Accurate estimation of tissue composition is important for risk assessment and for the selection of appropriate diagnostic modalities for detection. It may also have potential for improving computer aided diagnosis (CAD) performance by incorporating tissue composition information into CAD systems. Currently, while several computer algorithms have been developed to calculate breast tissue composition, they are generally based on a variety of subjective standards and do not necessarily provide quantitatively reliable results [5,10-14]. If such results are used in CAD system, their impact on CAD performance may be less than optimal.

A quantitative measure of breast tissue density, calibrated to some objective measure, would be of considerable value for both the estimation of cancer risk and for the development of CAD. Such a quantitative measure should be derived from more objective measures of 3D image data where tissue superposition does not obfuscate the correspondence between pixel values and tissue composition. We have previously developed a method to measure tissue composition from breast MR images. In this project we developed computer algorithms using physics based model to recover tissue composition information from 2D mammograms with compressed tissue and compared the results to measures from this MR data, as well as to radiologists' readings based on the BI-RAD standard [15]

Medical Imaging 2003: Image Processing, Milan Sonka, J. Michael Fitzpatrick, Editors,
Proceedings of SPIE Vol. 5032 (2003) © 2003 SPIE · 1605-7422/03/$15.00

2. METHODOLOGY

2.1 Mammography-based quantification of breast tissue composition

Digitized mammograms were used for this study. Prior to computing the amount of glandular tissue represented in these mammograms, the images were corrected for the non-linearity of radiographic density with respect to log energy exposure and for inhomogeneities of film exposure resulting from the heel effect and scatter. The images were also compensated for the non-uniformity of breast tissue thickness caused by breast compression. A physics based model was developed for calculating the fraction of glandular tissue at each pixel, and then applied to the normalized images. The total amount of glandular tissue was finally derived by integrating over all pixels within the breast area.

2.1.1 Correction for nonlinearity of film characteristic curve and digitizer

Radiographic density in mammograms has a non-linear relationship to log energy (X-ray) exposure, particularly at the high and low density extremes, and this would affect the accuracy of any composition estimates based on pixel values. To correct for this, it is necessary to linearize radiographic density in relation to log exposure. The mammograms used in this study were recorded on Kodak Min R 100 film, and a generic characteristic curve for this film was obtained from the manufacturer (Kodak, Indianapolis, IN). This characteristic curve was fitted with a polynomial spline function, which we used to convert pixel density values to log exposure values. Linearized density values were then derived by applying a linear mapping, designed to coincide with the central region of the film's characteristic curve, to convert log exposure values back to density values.

To correct for any non-linear relationship between optical density, as measured by the film digitizer, and log exposure, the digitizer was routinely calibrated to ensure linearity in the OD range of 0.1 ~ 3.8.

2.1.2 Correction for inhomogeneity of exposure intensity

Mammographic images are subject to exposure inhomogeneities resulting mainly from the heel effect and from scattered radiation. This may also affect the accuracy of quantitative estimates of composition. Once pixel values have been linearized as described above, we can partially correct for this inhomogeneity. The process begins by sampling pixel values within the image background, by finding the brightest pixel value within a kernel of 9x9 pixels through entire background area. These sampled points were then fitted by a second-order polynomial function

$$P(x, y) = a_0 + a_1 x + a_2 y + a_3 x^2 + a_4 y^2 + a_5 xy$$

using least squares methods, where x, y are pixel coordinate, a_i ($i = 0 \ldots 5$) are fitted coefficients, and $P(x,y)$ is the pixel value at coordinate of x, y. The distribution of background brightness was estimated using the fitted polynomial function. Each inhomogeneity corrected image was obtained by subtracting its calculated background brightness from the original image, pixel by pixel.

2.1.3 Correction for non-uniformity of breast tissue thickness

Besides non-linearity of pixel densitometry readings and inhomogeneity of exposure intensity, the non-uniformity of tissue thickness during breast compression may have the greatest impact on the association of pixel values with tissue types. We modeled compressed tissue thickness. and normalized images based on this model so that tissue thickness becomes uniform to eliminate thickness as a significant factor in estimating tissue composition from image density. A brief description of the method is presented here, and additional details can be found in previous publications [16-17].

Breast tissue thickness modeling -- Because of the elasticity of breast tissue, its thickness becomes tapered toward the skin line during compression. To model the thickness change, we first segmented the breast region from remainder of the image by employing algorithms previously developed for automatically detecting skin lines and pectoral muscles [18]. We then created a map of tissue thickness at each pixel as a function of the pixel's distance from the skin line. The distance between a pixel and the skin line was determined by an exhaustive search of the pixel's distance to each point on the skin line. Since the total X-ray attenuation depends on both the tissue attenuation coefficient and the tissue thickness, a measure of thickness can be estimated from projections of essentially pure tissue at each distance. It is difficult to directly measure X-ray attenuation. However, since pixel values had been linearized for log exposure, the

pixel value corresponding to pure tissue, such as fat, can ideally represent thickness change relative to its distance from the skin line. For each distance, the pixel value of pure fat tissue was approximated by taking the mean value plus one standard deviation from all of the pixels at that distance. The profile of the thickness values at all distance was then fitted to a curve function using exponential formula. This fitted thickness profile served as a tissue thickness model for the breast and was used for adjusting for breast tissue thickness.

Since our breast thickness model was based on approximation of pure fat tissue at each distance, it is likely that thickness was overestimated for extremely dense breasts where only few pixels corresponded to pure fat tissue. To correct the erroneous estimation caused by dense tissue, we examined the histogram of density values in breast tissue area prior to building thickness model. If the peak value was located within 20% of the dense end of the entire distribution range, the image was classified as dense tissue and the thickness function was fitted on only the first one third of the thickness profile, In particular, the center part of the tissue area on the image was not taken into account and was assumed to have constant thickness.

Breast tissue thickness correction -- The thickness coefficient was defined as the proportion of the maximum thickness and determined from the breast thickness model for each distance. The final step for correcting breast tissue thickness is to normalize breast thickness at each pixel by dividing the pixel value by the thickness coefficient at the distance of that pixel.

2.1.4 Quantitative measurement of Breast Tissue Composition

The method of quantitatively deriving tissue composition from adjusted density values was based on a model of the physics of X-ray attenuation, as tailored to mammography, which enabled us to estimate the fraction of glandular tissue, for each pixel, over the entire breast area.

After density values have been linearized as described above, the relationship between film density and exposure can be approximated as

$$D = C_o + kLog(E) \tag{1}$$

where D is the pixel value at an arbitrary pixel, C_o is the intercept of the pixel density-log exposure linear line, k is the film's gamma, and E is the exposure energy. Equation (1) can be rewritten as the difference of the background density of the film, and the total attenuation of the tissue projected onto a pixel:

$$D = D_0 - k\Sigma \mu_i x_i \tag{2}$$

where D_0 is density of film's background (i.e., non-attenuated exposure), μ_i is the attenuation coefficient of one type of tissue type, x_i is the thickness of that tissue type, and $\sum \mu_i x_i$ is the total attenuation of the tissue.

For this study, we assume a breast is basically composed of two types of tissue, fat and parenchyma, and most of the parenchyma is composed of glandular tissue. At an arbitrary pixel, if we define x_{fat} as the relative thickness of fat tissue and $(1 - x_{fat})$ as the relative thickness of parenchymal tissue, then equation (2) can be expended to

$$D = D_0 - k[x_{fat}\mu_{fat} + (1 - x_{fat})\mu_{parenchyma}] \tag{3}$$

If $D_{(fat)}$ is the pixel density corresponding to pure fat, then x_{fat} becomes 1 and with equation (3) we have

$$D_{(fat)} = D_0 - k\mu_{fat}$$

or

$$\mu_{fat} = \frac{D_0 - D_{(fat)}}{k} = A$$

We also define $R = \mu_{parenchyma} / \mu_{fat}$.

By rearrangement, equation (3) can be rewritten as

$$x_{fat}(\mu_{fat} - \mu_{parenchyma}) + \mu_{parenchyma} = \frac{D_0 - D}{k} = B$$

To estimate relative thickness of fat tissue at an arbitrary pixel location, we derived a formula from the above equation

$$x_{fat} = \frac{B - \mu_{parenchyma}}{\mu_{fat} - \mu_{parenchyma}} = \frac{B - RA}{A - RA} = \frac{\frac{B}{A} - R}{1 - R} = \frac{\left(\frac{D_0 - D}{D_0 - D_{(fat)}}\right) - R}{1 - R} \qquad (4)$$

The relative thickness of parenchymal tissue at a pixel is expressed as $1 - x_{fat}$. The background density was measured as the mean pixel value in a 15 × 15 pixel area with no object present. D_{fat} was calculated as the mean value of the 15% of pixels having the highest values within the breast area. The ratio of the breast parenchyma attenuation coefficient to the fat attenuation coefficient, R, was obtained from published literatures [19-20]. The final tissue composition measure for each image was the proportion of parenchymal tissue in the entire breast area, which was obtained by dividing the sum of the relative thickness of all parenchymal tissue pixels by the total number of pixels.

2.2 Breast tissue composition measures from MRI data

In order to process a large number of MR images, we developed an automated method to estimate breast composition from axial MR images. The images were T1 weighted gradient echo acquisitions without fat suppression, acquired on a 0.5 T Aurora dedicated breast scanner. We first identified an artifact free ROI within the air region of the image. To ensure exclusion of artifacts, the ratio of the mean (μ) to the standard deviation (σ) within the region was tested to match the expected value of 1.91 following Henkeleman [21]. A binary mask was generated by setting a threshold of $\mu+N\sigma$. Holes in the mask were filled using a morphological closing operation. Intensity inhomogeneities were reduced by applying homomorphic filtering. The image was blurred with a wide Gaussian filter. The original image was then divided by this blurred image. The chest location was estimated by the first nonzero point at the midpoint between the breasts. A straight cut in the image was made perpendicular to the breasts at this point to separate the breasts from the chest. A crisp clustering algorithm was then used to segment the masked image into two compartments (parenchyma and fat). The threshold level and structuring elements for the morphological closing were optimized based on 12,000 randomly selected points taken from eight consecutive patients, which were classified by an observer. The optimal threshold level was found to be $\mu+27\sigma$. The morphological closing was achieved by three successive dilations with a cross like structuring element (i.e., ones along the principle axis, zero elsewhere) of size 9x9x7 followed by three successive erosions with the same structuring element. In selected review of cases the algorithm was seen to perform quite well. Exceptions included very large, fatty breasts where the proximity of the breasts to the receive coils resulted in large inhomogeneities which were not adequately corrected. Also in some patients ghosting of the heart resulted in inclusion of relatively large air regions between the breasts into the mask. Both of these errors resulted in an overestimation of breast parenchyma.

2.3 Breast tissue composition assessment from radiologists

Subjective scores were determined by mammographers' readings from digitized mammograms using the BI-RAD standard. Digitized images, with resolutions of 50×50 μm^2 and 12 bits, were read on a 1024×2048 CRT monitor. Three experienced mammographers individually scored all of the cases in the study's database, and their scores were averaged

to get the mean values reported as final scores. For each breast, there were separate scores for the cranial-caudal views (CC view) and the medial-lateral-oblique view (MLO view), and a score for the breast obtained by viewing both the CC and MLO views simultaneously. When individual views (including CC views and MLO views) were presented to mammographers for scoring, images from the cases were randomly mixed so that they were not identifiable for pairing of consecutive images. The order of image pairs (CC view and MLO view) of breasts for scoring was also randomized so that consecutive pairs from the same patient were unlikely.

Mammographers' scores for composition were based on the BI-RAD standard, which coarsely classifies tissue composition into four categories – category I indicates that the breast is almost entirely fat; category II indicates that there are scattered fibroglandular densities; category III indicates that the breast is heterogeneously dense; and, category IV indicates that the breast tissue is extremely dense. To accommodate patterns between categories, and variations within a category, radiologists reported on a 40 unit scale, which had been defined such that units of 1-10 represents BI-RAD category I, 11-20 represents BI-RAD category II, 21-30 represents BI-RAD category III and 31-40 represents BI-RAD category IV.

2.4 Correlation analysis

To compare the results from the computerized measurement of digitized mammograms with scores from MR images, or with scores from mammographers' readings, least square distances were calculated to obtain correlation coefficient for the two sets of data.

Each MRI score from one breast was paired respectively with mammography-based score from CC view, from MLO view or from mean value of CC view and MLO view of corresponding mammogram. To compare scores from mammographers' readings, the scores of CC view, MLO view, and CC view together with MLO view from mammographers' readings were paired with mammography-based scores of CC view, MLO view, and mean value of CC view and MLO view, respectively.

2.5 Image database

We collected 145 cases for this study. Each case had two views (CC view and MLO view) of mammograms and a corresponding MR exam. For each case, the time difference between the acquisitions of the two image sets was less than six months.

Mammograms were digitized with a Lumisys digitizer (Sunnyvale, CA). Digitized mammograms had a spatial resolution of 50×50 μm^2 and 12-bit brightness resolution. For this particular study, which evaluated only global characteristics of images, the initial spatial resolution was higher than what was required, and was reduced to 400×400 μm^2 by filtration. The digitizer was routinely calibrated to ensure linear translation from radiographic intensity of film to optical density over the range of 0.1 to 3.8 optical density.

MR images were acquired on a 0.5T Aurora dedicated breast scanner (Aurora Imaging Technology, North Andover, MA). The Aurora breast coil was used. Images were acquired on axial plane using a 3D (2.5mm slices) gradient echo sequence. To keep the processing simple, in particular to avoid registering breast images, we limited ourselves to a single spectral band (T1 without fat suppression).

3. RESULTS

Table 1 lists the linear correlation coefficients for view-based and breast-based comparisons between computer scores derived from mammograms and the scores from MR images (A), and between computer scores from mammograms and the scores from mammographers' readings (B). The results indicate that scores between MR images and mammography-based computations have higher linear correlation than the scores between mammographers' readings and mammography-based computations. Scores based on two-views generally gives better correlation than those based on only one view.

When the correlation data is plotted graphically (Figure 1), it appears that the relationship between the scores from mammographers' readings and scores from mammography-based computation was not linear, suggesting that qualitative estimate and quantitative measures can not be directly mapped to each other.

	Mammography-based Computer Scores		
	CC view	MLO view	Two views (CC view and MLO view)
MRI Scores	0.62	0.50	0.65
Radiologists' Scores	0.57	0.51	0.61

Table 1: Linear correlation coefficients of breast tissue composition measurements between mammography-based computer scores and MRI scores, and between mammography-based scores and radiologists' scores.

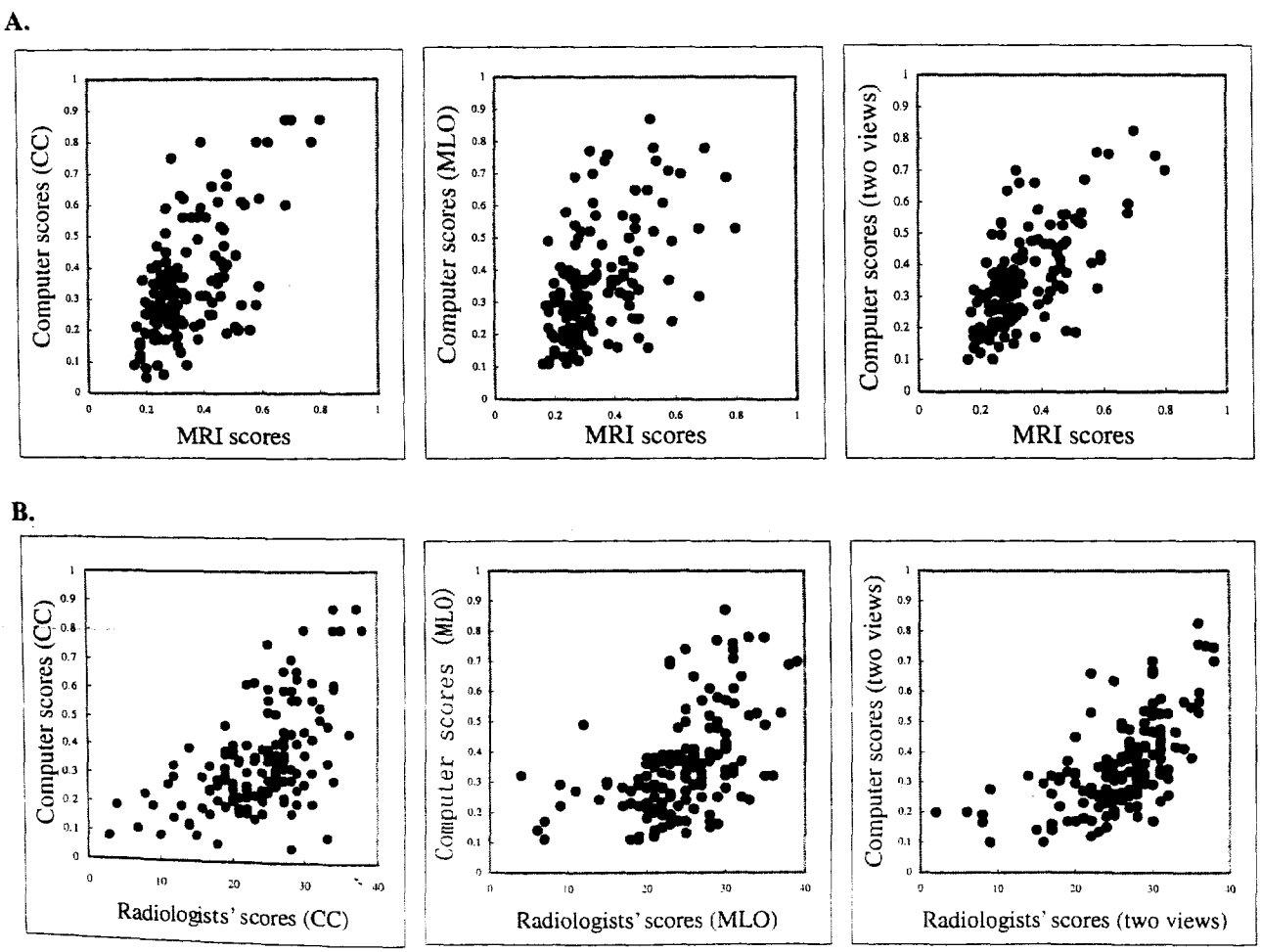


Figure 1. Scatter plots of mammography-based computer scores versus breast tissue composition measures from MRI data (panel A), and mammography-based computer scores versus radiologists' scores (panel B).

4. DISCUSSION

We have shown in this study that it is feasible to recover information of breast tissue composition from 2D mammographic data, despite tissue compression, by using computerized algorithms developed in this project. Furthermore, objective measures from 3D MRI data can be a viable standard for tissue composition estimates, as suggested by the high linear correlation between MRI scores and mammography-based computer scores, shown in our results. Although the results look promising, they are not highly optimal. I major limitation of this work is that the MR image data used in this study was less than optimal, being acquired on a 0.5T system. This caused some difficulty in removing artifact from the MR data and in adjusting parameters associated with segmentation and background correction in the MR images. Two major problems with this set of MR image data, as mentioned in the methods section, were inhomogeneity of image background and ghosting effect, which tend to overestimate amount of parenchymal tissue in the breast region. By getting better data, we would expect more reliable scores from MR images.

Several steps were involved in mammography-based tissue composition computation. The crucial one among those steps is thickness correction, which has limitations on its application to certain types of mammograms. When breast tissue is very dense, the estimated thickness model can be incorrect because deriving a thickness profile from the image data assumes that regions of pure fat tissue exist, and its attenuation can be approximated from statistical values of the image data. Using a histogram of the distribution of pixel densities, to identify very dense tissue and adjust its thickness profile accordingly, improved estimates for most of the dense breast images. Another factor that can cause incorrect thickness estimation is the existence of a skin flap near the chest wall, or lumpy breast tissue caused by the reduction in fat tissue with age. Images with these features cause locally irregular thickness changes, and the thickness model we developed in this project does not provide an adequate mechanism to accommodate those images.

Quantitative measures of breast tissue composition become particularly desirable for incorporation into CAD systems. Currently there is no accepted standard that can be adopted as a gold standard for composition, and subjective and qualitative methods result in large variations, and possibly non-linear relationships to actual composition, as indicated in our results. This study suggest that quantitative measures derived from 3D MR image data may provide a good standard for quantitative computation of breast composition.

Even though many improvements are needed to our algorithms, and in the MR data we use for comparison, the preliminary results presented in this study suggest that breast MR provides a promising standard for composition estimation, and that the improvements in deriving composition from mammograms will likely enhance our ability to accurately determine the breast cancer risk associated with breast density.

ACKNOWLEDGEMENT

This work is supported in part by grant IMG2000362 from the Susan G. Komen Foundation, grants CA80836 and CA77850 from the National Cancer Institute, National Institutes of Health and also by the US Army Medical Research Acquisition Center, 820 Chandler Street, Fort Detrick, MD 21702-5014 under Contract DAMD17-02-1-0549. The content of the information contained here does not necessarily reflect the position or the policy of the government, and no official endorsement should be inferred.

REFERENCES

1. J.N. Wolfe. A.F. Saftlas. M. Salene. "Evaluation of mammographic densities: A case-control study," *AJR* 148:1087-1092 (1987).
2. A.F. Saftlas. J.N. Wolfe. R.N. Hoover. et al.. "Mammographic parenchymal patterns as indicators of breast cancer risk," *Am J Epidemiol* 129:518-526 (1989).

3. N.F. Boyd, B. O'Sullivan, J.E. Campbell, et al., "Mammographic signs as risk factors for breast cancer," *Br J Cancer* 45:185-193 (1982).
4. V.L. Ernster, S.T. Sacks, C.A. Peterson, "Mammographic parenchymal and risk factors for breast cancer," *Radiology* 134:617-620 (1980).
5. N.F. Boyd, J.W. Byng, R.A. Jong, et al., "Quantitative classification of mammographic densities and breast cancer risk: results from the Canadian National Breast Screening Study," *J Natl Cancer Inst* 87:670-675 (1995).
6. C.H. van Gils, J.D. Otten, A.L. Verbeek, J.H. Hendriks, "Mammographic breast density and risk of breast cancer: Masking bias or causality?" *Eru J Epidemiol* 14:315-320 (1998).
7. R.L. Egan, R.C. Mosteller, "Breast cancer mammography patterns," *Cancer* 40:2087-2090 (1977).
8. B. Threatt, J.M. Norbeck, N.S. Ullman, R. Kummer, P. Roselle, "Association between mammographic parenchymal pattern classification and incidence of breast cancer," *Cancer* 45:2550-2556 (1980).
9. N.F. Boyd, G.A. Lockwook, J.W. Byng, D.L. Tritchler, M.J. Yaffe, "Mammographic densities and breast cancer risk," *Cancer Epidemiol Biomarkers Prev* 7:1133-1144 (1998).
10. C. Zhou, H.P. Chan, N. Petrick, et al., "Computerized image analysis: Estimation of breast density on mammograms," *Med Phys* 28:1056-1069 (2001).
11. E. Sala, R. Warren, J. McCann, S. Duffy, R. Luben, N. Day, "Mammographic parenchymal patterns and breast cancer natural history," *Acta Onclogica* 40:461-465 (2001).
12. N. Karssemeijer, "Automated classification of parenchymal patterns in mammograms," *Phys Med Biol* 43:365-378 (1998).
13. J.W. Byng, N.F. Boyd, E. Fishell, R.A. Jong, M.J. Yaffe, "Automated analysis of mammographic densities," *Phys Bed Biol* 41:909-923 (1996).
14. P.G. Tahoces, J. Correa, M. Souto, L. Gómez, JJ.Vidal, "Computer-assisted diagnosis: the classification of mammographic breast parenchymal patterns," *Phys Med Biol* 40:103-117 (1995).
15. American College of Radiology (ACR). *Illustrated breast imaging reporting and data system (BI-RADS™)*, 3rd ed. Reston VA: American College of Radiology, 167-176 (1998).
16. X.H. Wang, B. Zheng, Y-H. Chang, W.F. Good, "Correction of digitized mammograms to enhance soft display and tissue composition measurement," *Proc. SPIE* 4319:328-336 2001).
17. X.H. Wang, W.F. Good, B.E. Chapman, Y.H. Chang, W.R. Poller, T.S. Chang, L.A. Hardesty, "Automated breast tissue composition assessment with tissue thickness corrected mammograms," *Am J Roentgenol AJR* 180:257-262 (2003).
18. W. F. Good, B. Zheng, Y. H. Chang, X. H. Wang, and G. Maitz, "Generalized Procrustean deformation for subtraction of mammograms," *Proc. SPIE* 3661:1562-1573 (1999).
19. P.C. Johns, M. J. Yaffe, "X-ray characterization of normal and neoplastic breast tissue," *Phys in Med Biol* 32: 675-695 (1987).
20. R. P. Highnam, J. M. Brady, B. J. Shepstone, "Computing the scatter component of mammographic images," *IEEE Trans med. Imag* 13: 301-313 (1994).
21. R.M. Henkelman, "Measurement of signal intensities in the presence of noise in MR images," *Med. Phys.* 12:232-233 (1985).

Appendix 3

Feature-based differences between mammograms

Walter F. Good, Xiao Hui Wang, Glenn Maitz
Department of Radiology, University of Pittsburgh, Pittsburgh, PA 15261

ABSTRACT

A novel technique for assessing local and global differences between mammographic images was developed. This method uses correlations between abstract features extracted from corresponding views to compare image properties without resorting to processes that depend on exact geometrical congruence, such as image subtraction, which have a tendency to produce excessive artifact. The method begins by normalizing both digitized mammograms, after which a series of global and local feature filters are applied to each image. Each filter calculates values characterizing a particular property of the given image, and these values, for each property of interest are arranged in a feature vector. Corresponding elements in the two feature vectors are combined to produce a difference vector that indicates the change in the particular properties between images. Features are selected which are expected to be relatively invariant with respect to breast compression.

Keywords: Mammography, Computer-aided detection, Breast symmetry

1. INTRODUCTION

In evaluating mammographic studies, mammographers frequently need to compare images – either to determine the degree of contralateral symmetry or to determine changes over time for a fixed view. Traditionally, mammographers assess the degree of asymmetry between corresponding views of left and right breasts, or changes between a current study and a prior study, as indicators of the likelihood of the presence of abnormalities. In an effort to mimic the performance of mammographers, CAD algorithms sometimes attempt a rudimentary measure of differences between mammograms by subtracting images after they have been placed in registration, to the extent possible, by rigid translations and rotations[1]. Because of compression induced geometric deformations, and ambiguities introduced by the superposition of tissues, both of which are characteristic of projection mammography, the rigid registration and subtraction process generates difference images which contain considerable artifact and are visually unappealing. Thus, these methods are an inefficient mechanism for detecting the kinds of difference information of value to mammographers. The real problem is that the similarities and differences of interest actually relate to features and abstract objects and topological properties, rather than to geometric equivalence.

1.1 Information content of image differences

There is abundant evidence that difference images in certain procedures can contain diagnostically useful information, despite the presence of substantial subtraction artifact. In the diagnosis of chest radiography, it has been demonstrated that adding temporally subtracted images to the diagnostic process can, in many situations, significantly improve diagnostic accuracy and reduce mean interpretation time[2]. Bilateral subtraction images, which are frequently used in CAD algorithms, have been shown to contain diagnostic information[1,3-6]. However, subtraction of mammograms is a difficult task due to inconsistencies in breast compression, skewing of the three-dimensional breast relative to the image projection during compression and lack of easy landmarks on the breast that can be used to facilitate optimal image registration. Previous attempts to develop acceptable subtraction methods for mammograms, including methods for both bilateral subtraction[1,3-5] and temporal subtraction[7-9], have had only partial success. As a result, the subtraction of mammograms has not been employed clinically. Furthermore, the inability of current methods to put mammographic images in accurate registration for subtraction, or to use other methods to find differences, greatly impairs the ability of CAD to exploit this kind of information.

Medical Imaging 2003: Image Processing, Milan Sonka, J. Michael Fitzpatrick, Editors,
Proceedings of SPIE Vol. 5032 (2003) © 2003 SPIE · 1605-7422/03/$15.00

1.2 Overview of our approach

At present there are no algorithms that successfully capture the essence of image differences, as seen by mammographers. In our first effort to find an alternative to image subtraction, we have developed a method which attempts to perform a feature-based analysis of differences between images. This has involved the integration of several technologies to provide an efficient mechanism for performing quantitative analysis of mammographic images. A few of these methods are briefly described here, and elaborated on in the methods section below.

1.2.1 Adjustment for tissue thickness variations during breast compression

Techniques for automatically making the appropriate local image corrections, based on the variations in tissue thickness during breast compression after correcting for the nonlinearities of the characteristic curves of film and digitizer, have previously been developed by us[10] and by others[11-13]. Application of these methods can improve the soft display of mammograms as well increase the accuracy of quantitative estimates of tissue composition.

1.2.2 Measure of tissue composition

We have recently developed a computerized method for calculating tissue composition from digital mammographic data, which has been calibrated against estimates derived from breast MRI[10]. By first correcting for the nonlinearity of the detectors' characteristic curve and for the thickness of the breast toward the skin line during compression, we have been able to implement a classification mechanism that provides accurate estimates of composition.

1.2.3 Geometric deformation

A simple computerized method for deforming mammograms has been developed, which puts nipples, chest walls and skin lines in registration, and which we believe is nearly optimal with respect to the tradeoff between registration accuracy and loss of diagnostic information[14,15]. The transformation is plausible in the sense that it is invariant with respect to translation, rotation, and degree of magnification or minification, and it behaves geometrically in a manner that counteracts certain kinds of differences resulting from breast compression.

1.2.4 Multi-image CAD

Certain kinds of features related to focal densities (e.g., location, texture, degree of spiculation, and integrated density difference) tend to be relatively invariant, or at least predictable, with respect to breast compression and projection during mammography. These features show a strong correlation between views, for densities that are projections of a compressed mass, but are independent for unrelated densities. Methods have been developed to extract considerable information about focal densities from such correlations, or lack thereof[16].

1.2.5 Walsh-Hadamard Transform adapted for mammographic images

Because breast tissue is projected onto only part of a mammographic film, when traditional two-dimensional transforms are applied to mammographic images, the image background usually has a significant, and undesirable, impact on the result. Furthermore, most of these transforms are effectively based on a rectangular decomposition of the image area, and this has limited relevance to the tissue patterns as presented in projection mammography. In order to more accurately characterize the tissue patterns in mammograms, and to attempt to retain some relevance to actual fibroglandular tissue structure, we developed a transform mechanism specific to mammography. This method is basically an adaptation of the Walsh-Hadamard transform[17], incorporating a non-rectangular decomposition, which has been constrained to the image area representing tissue.

1.3 Hypothesis and aims

The goal of this project was to develop an automated technique for assessing local and global differences between images, without resorting to processes that depend on local geometrical registration. This was inspired by the hypothesis that corresponding contralateral mammographic views, and temporal sequences of mammograms, contain useful diagnostic information that is independent of the information that can be derived by analyzing the images separately, and that the bulk of this cannot be extracted by algorithms based on image registration and subtraction. Furthermore, we assume that, by employing a symbolic as opposed to a geometric methodology, correlations between abstract features extracted from corresponding views can be used to match image components without relying on, or expecting, exact geometrical congruence.

A question arises as to what kind of differences are actually of clinical significance. Differences reported by mammographers usually relate to changes is some specific set of image properties, whereas differences in individual pixel values, in isolation, are relatively meaningless. In this initial investigation we focused primarily on features related to breast composition, distribution patterns of fibroglandular tissue, and on certain aspects of image texture.

2. METHODS

2.1 Overview of technique

This project was a preliminary attempt to integrate various techniques for computerized analysis of mammograms, for the purpose of finding differences between similar images without resorting to processes that principally involve image registration and subtraction. The specific process we are studying begins by correcting both mammograms for the combined characteristic curve of film and digitizer (in the case of film images) to produce pixel values proportional to exposure, and then locally adjusts for nonuniformities in the thickness of the breast during compression. We then geometrically deform the images to match a semicircular template by using a nonlinear geometric transformation we have previously published, which places nipples, chest walls and skin lines in accurate registration to the template. Tissue densities are corrected locally by the Jacobian of this transformation and then globally so that pure fat voxels and pure fibroglandular tissue voxels (if such voxels actually existed in the mammogram) are mapped onto standardized values, by using methods we have developed for identifying tissue composition of specific voxels.

Figures 1A and 1D show CC views of the left and right breasts of a woman having a malignancy in the left breast. The thickness-corrected versions of these images are shown in Figures 1B and 1E. The images depicted in Figures 1C and 1F have been geometrically transformed to match a semicircular template. It is the images in Figures 1C and 1F to which further analysis will be applied.

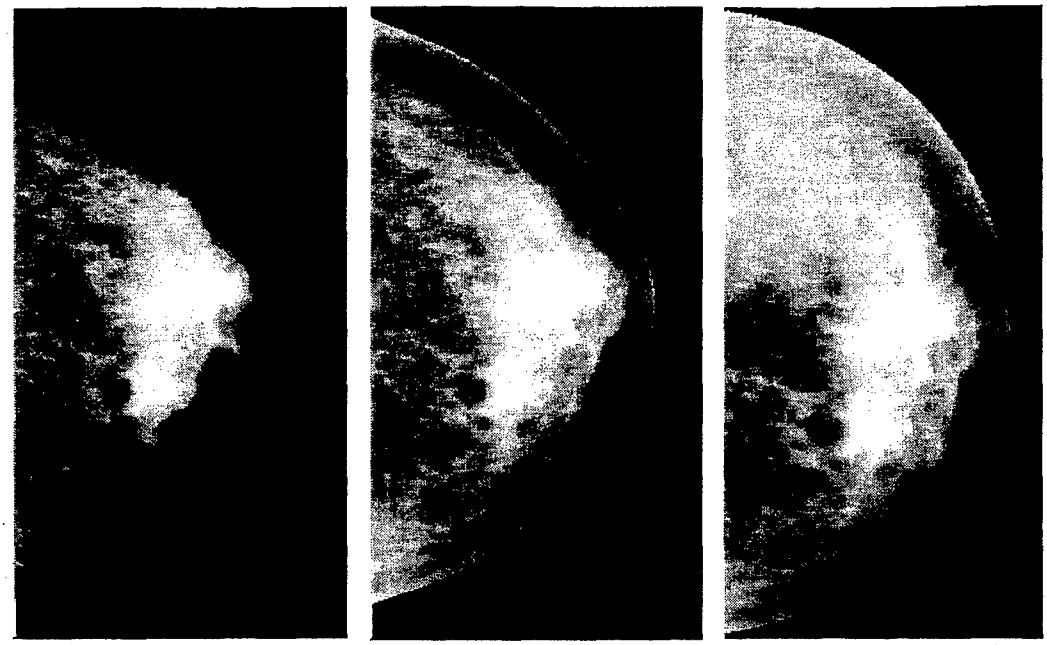
Once images are in registration with the template, and their gray scales have been adjusted, we apply a series of feature filters to characterize the image properties of interest. A variety of global features, including measures of tissue composition, slope of logarithm of power spectra and image texture, can be made largely invariant to breast compression and imaging parameters and are useful for characterizing breast images. Additional features for categorizing the pattern of fibroglandular tissue are derived by applying a modified version of the Walsh-Hadamard transform that has been adapted specifically for evaluating tissue patterns in mammography. Features are placed in a feature vector that represents the global properties of the image that are of interest in the particular context. Appropriate comparison of feature vectors reveals differences between image properties. These methods represent a novel approach to evaluating differences, and reflect more accurately the process performed by mammographers in making such judgments.

Our method also provides a mechanism for analyzing changes in focal densities by employing multi-view CAD algorithms, which we previously developed for analyzing ipsilateral pairs, to extract and compare features related to focal densities from both images. This process identifies objects, and their features, in each view, and evaluates all possible pairs of objects (with one from each view) to construct the most likely mapping between objects within the two images. Examples of the kind of spatial features we derive for focal densities, that are appropriate for multi-image CAD are: radial distance from nipple; length projection of density onto the NAL; integrated contrast difference relative to background; complexity of boundary around a density; local image texture; and, margin definition of a density. The mechanism for comparing focal densities between ipsilateral image pairs has been previously published and is not presented here[16].

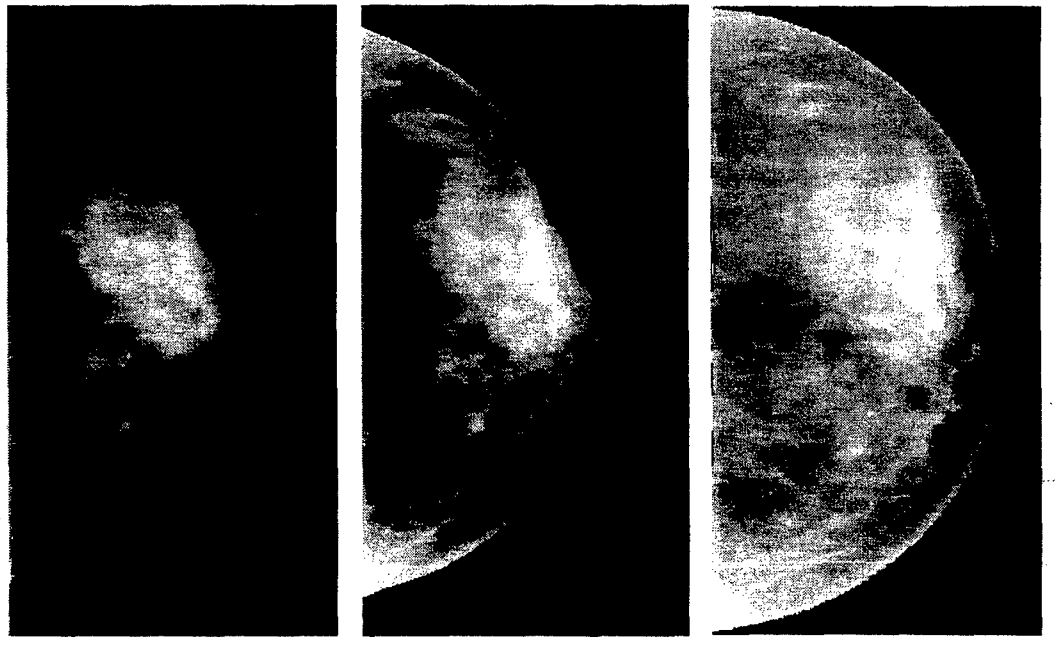
In this preliminary study, we limited our attention to detecting differences in the distribution of fibroglandular tissue. Other kinds of global features could be handled in a similar manner.

2.2 Individual image normalization

In order to facilitate the quantitative comparison of images, it is desirable to correct the grayscales of images for nonlinearity of characteristic curves of detectors[10], to provide values linearly proportional to logarithm of exposure, and then to correct image densities for non-uniformities in the thickness of breast tissue during breast compression.



A B C



D E F

Figure 1 CC views of the left and right breasts of a woman having a malignant mass in the left breast. Original images are shown in A and D. Thickness-corrected and geometrically transformed versions are shown in B and E and in C and F, respectively.

2.2.1 Correction for characteristic curve of digitizer and film

The characteristic curve for each particular film is obtained from the manufacturer. Our routine quality assurance process on the digitizer maintains a linear relationship between pixel value and optical density. The characteristic curve correction algorithm is written to use the inverse of the digitizer characteristic curve to convert pixel values to optical density values, and then us the inverse of the film's characteristic curve to convert optical density to film exposure.

2.2.2 Adjusting pixel values for tissue thickness

Although there are computationally more efficient methods, we measure the distance of each pixel to the skin line with a simple exhaustive search. We divide the range of possible distances into a small number of intervals, and the mean and standard deviation are calculated for pixels within each interval. An appropriate function is fitted, with constraints, to the means plus one standard deviation. This "correction" function represents the change in pixel value with respect to distance from the skin line, and actually indicates tissue thickness relative to this distance. Figure 2 shows the correction function for the image in Figure 1A. For each pixel, the correction function is used to calculate a correction value from the pixel's distance value, and this correction value is used to adjust the pixel value. In the central regions of the breast area, the correction values are 1 and these pixels are left unchanged[10].

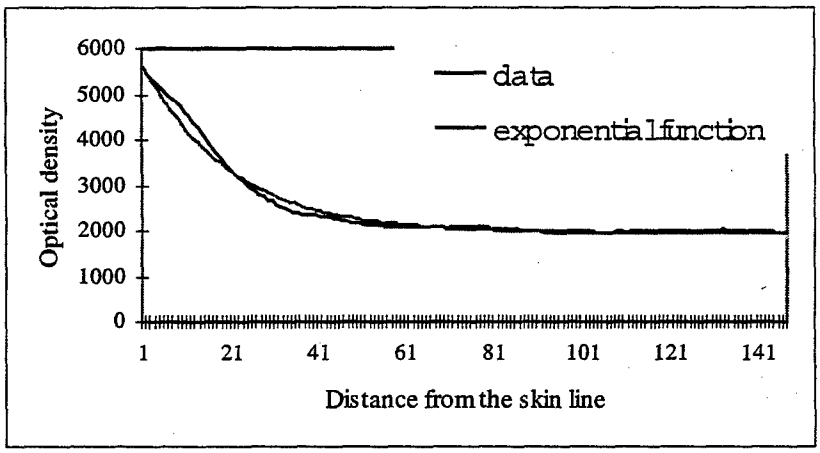


Figure 2 Thickness correction curve, and fitted curve, for image shown in Figure 1A. The curves are approximately inversely proportional to relative breast thickness. Distance is measured in pixels.

2.2.3 Estimating tissue composition

After making the above adjustments to pixel values, at each point in a view we estimate % fibroglandular tissue, by deriving certain features directly from a local histogram and using a neural network, trained with values from breast MRI, to obtain a local estimate of composition. A global composition value is determined for the image by integrating over the tissue area[10]. This enables us to adjust grayscales to an absolute range – so that pixels corresponding to pure fat and to pure fibroglandular tissue, if they existed, would always fall at fixed pixel values.

2.3 Geometric Transformation of images

To facilitate a comparison between mammograms, normalized images are deformed to match a template, which is defined to be a semicircular region, with the nipple location specified as being the middle of the arc of the semicircle. The transformation is designed to map the mammogram's nipple onto the nipple location on the template, and map the skin line onto the semicircular boundary of the template. The total deformation of the transformation is measured by:

$$M(I) = \frac{1}{Template\ Area}\sqrt{\sum_r \sum_c \left[\det J\left(\mathcal{J}\left(I_{r,c}\right)\right)\right]^2},$$

where, $J\left(\mathcal{J}\left(I_{r,c}\right)\right)$ is the Jacobian of the transformation of the source image *I*, at coordinate (r,c),

and the resulting value is included in the global feature vector as an additional feature characterizing breast shape. A fully automatic nonlinear geometric transformation capable of performing such a deformation has been previously published[14,15], but was originally designed to deform one mammographic image to the shape of a second mammographic image, putting their skin lines, chest walls, and nipples in exact registration for the purpose of image subtraction. The adaptation of the transformation for this project is described here.

The transformation, which we refer to as the relative radius polar coordinate method, is diagramed in Figure 3. The technique begins by automatically detecting, on the input image, the pectoral muscle if it is an MLO view (e.g., ***OL*** in Figure 3), the skin line and nipple location, ***N*** in Figure 3. The skin line is described by a spline function, to facilitate the processing described below. A polar coordinate system is established with the origin, ***O,*** at the intersection of the nipple axes line (NAL), ***ON,*** with the line indicating the pectoral muscle, on an MLO view or with a line parallel to, and *0.5 cm* beyond, the edge of the film on a CC view. A corresponding polar coordinate system is established on the template with its origin, ***O',*** as indicated in Figure 3.

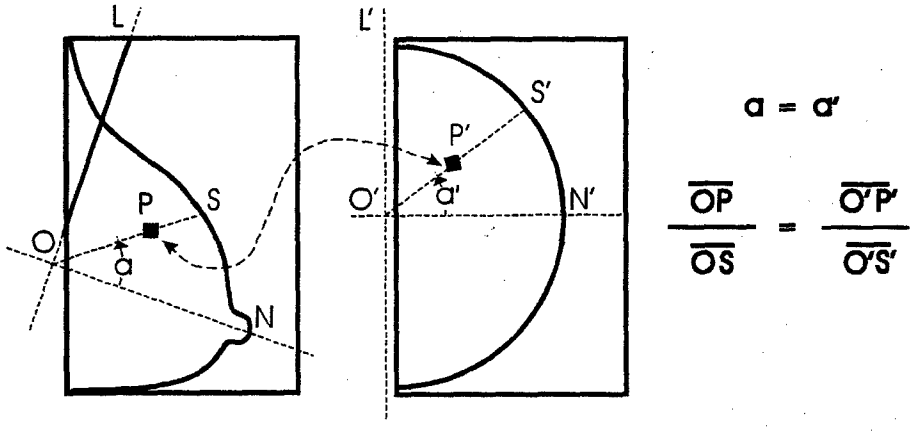


Figure 3 Schematic representation of geometric transformation.

Tissue pixels within a mammogram are identified by a relative polar coordinate, which we define to be the angle, **α**, of their position vector relative to the NAL and their fractional distance between the origin and the skin line. For each pixel ***P'*** in the reference template (e.g., right image in Figure 3), a point, ***P,*** having the same relative polar coordinate, is found in the image to be deformed. The pixel at location ***P'*** in the template is then given the value of the point at ***P***, as determined by the nearest pixel or through interpolation. In general, this mechanism allows two mammograms to be deformed so that their pectoral lines, NALs and skin lines are all in complete registration. In this case, the source mammogram is deformed such that its skin line, pectoral muscle and nipple are in registration with the corresponding features on the template. As an image is transformed to the reference geometry, pixel values are adjusted for the local expansion factor of the transformation by dividing each pixel value by the local Jacobian of the transformation at the pixel's position.

2.4 Walsh-Hadamard transformation adapted for mammographic images

Tissue patterns in mammograms can be characterized by classifying combinations of coefficients from transformations such as the Fourier transformation or the Walsh-Hadamard transformation. For this project we developed a version of

the Walsh-Hadamard transformation that was constrained to operate on only the image area representing breast tissue. The Walsh-Hadamard transform decomposes signals by applying an orthonormal set of rectangular waveforms (basis vectors). The individual waveforms are not generally periodic, but they can be ordered by a property referred to as sequency, which is the number of sign changes in a waveform[18]. The two dimensional version of the transform is separable and there exist efficient algorithms for it computation. To constrain the transformation to tissue areas, we employ the geometric transformation described above to deform the tissue area of a mammogram to coincide with the semicircular template. The semicircular area is divided horizontally by chords that partition it into regions of equal area, and divided vertically by lines originating at the nipple position and ending at the chest wall side of the template as depicted in Figures 4A and 4B respectively, for the 8×8 case. These two patterns are superimposed to form the basic grid upon which the transform is calculated. The ordered form of the transform can be calculated as:

$$\mathcal{W}(r,s) = \frac{1}{N}\sum_{i=0}^{N-1}\sum_{j=0}^{N-1} I(i,j)(-1)^{p(i,j,r,s)}$$

where,

$$p(i,j,r,s) = \sum_{k=0}^{N-1}\left(h_k(r)i_k + h_k(s)j_k\right)$$

and

$$\text{for } k \neq o: \quad h_k(r) = r_{N-k} + r_{N-k-1}$$

$$h_0(r) = r_{N-1}\,.$$

Figure 5 depicts the set of basis functions for the 8×8 case.

While this method has demonstrated some clear advantages over traditional transform techniques, it has certain significant limitations. For characterizing tissue patterns, at a scale that is appropriate for diagnostic purposes, an 8×8 or 16×16 transform may be adequate. However, due to the complexity of partitioning a discrete semicircle into ever-finer regions of equal area, the transformation cannot be extended to resolutions on the order of the pixel size. While this problem can be overcome by defining the template to be sufficiently large, or by adopting an elaborate interpolation scheme, the impracticality of these processes limits its usefulness. Also, the use of a non-rectangular decomposition of the image area greatly increases the difficulty in implementing fast algorithms for the calculation, and partially neutralizes the computational efficiency of the Walsh-Hadamard transform.

2.5 Comparison of Feature Vectors

Changes in individual features between two images are determined by comparing the components or their respective feature vectors. If a given property is similar between images, then the corresponding features in the feature vectors should be nearly equal. In general, we define the difference between a feature in two feature vectors to be:

$$D_i(I,J) = \left[\log F_i(I) - \log F_i(J)\right]^2,$$

where, F_i is the filter that measures feature i. Changes in image properties, that involve multiple features, are calculated as a weighted sum of these differences over all features related to the property.

$$Property_k = \sum_{i \in F_{\text{Property}}} w_i D_i\,.$$

The associations between properties and subsets of features will eventually be implemented as a set of neural networks to allow for nonlinear summations of feature differences, but this will require training the neural networks on a large set of training cases.

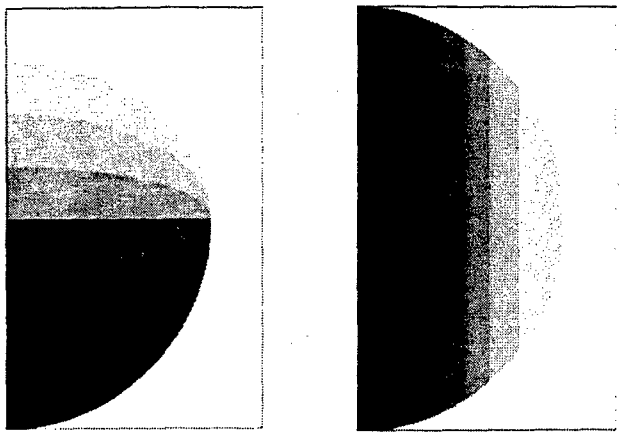


Figure 4 Horizontal and vertical partitioning of semicircle template, for the mammography-adapted Walsh-Hadamard transform (N=8).

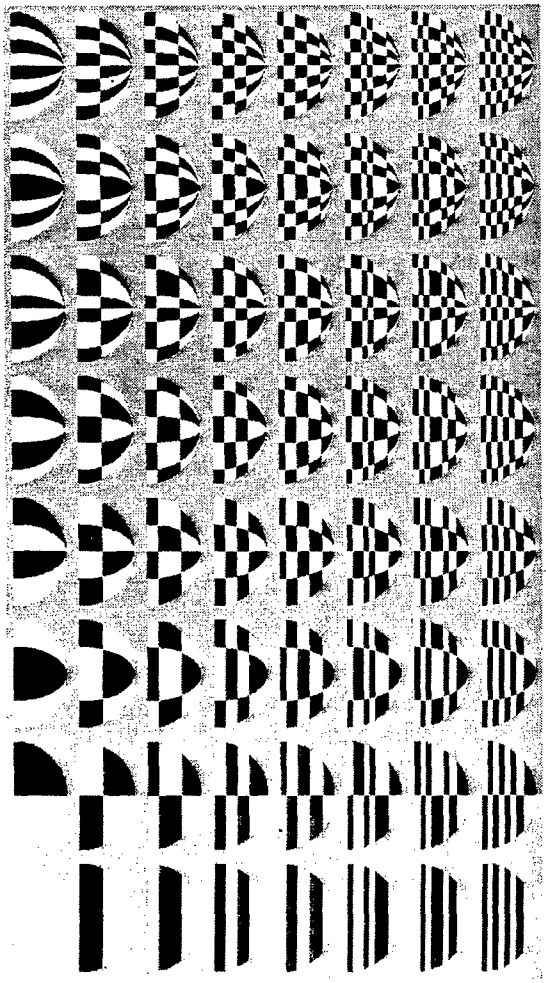


Figure 5 Set of two-dimensional basis vectors (N=8) for the mammography-adapted Walsh-Hadamard transformation. (1-White, 0-Black)

2.6 Preliminary Evaluation

To evaluate the feasibility of the methods described herein, we calculated global feature vectors for a set of 100 images, which included ipsilateral views, contralateral views, and current and prior studies. Cases were selected from a pool of routine mammographic examinations performed at the University of Pittsburgh Medical Center, and were digitized using a Lumisys-150 (Lumisys, Sunnyvale, CA) high-resolution, high-contrast sensitivity laser film digitizer that produces a scan matrix of 4,096 x 5,120 pixels for an 8 x 10-inch film, by digitizing at a 50-μm sampling interval. While this project focused on the use of digitized films, most of the effort will be applicable and/or adaptable to images acquired using direct digital systems.

We have applied this method to measuring changes in a variety of image properties. Two representative studies are described below and in the results section.

2.7 Measure of similarity in tissue pattern

A summary measure of similarity between two images, *I* and *J*, was defined to be:

$$\mathrm{P}(I,J)=\sum_i w_i\left[\log P_i(I)-\log P_i(J)\right]^2,$$

where P_i is pattern filter i, and w_i is its weighting factor. Based on the various combinations of images contained in the case set, we evaluated the above measure for corresponding left and right contralateral views and for ipsilateral pairs, for all cases where the images were available. To determine the difference in this measure between image pairs acquired from one woman or one breast, and pairs of images selected randomly from the pool of images, we performed the measure on 50 pairs of CC views that were acquired on different women. We repeated the evaluation for each case where we had a current image and a corresponding prior view from the same woman.

2.8 Change in tissue pattern between current and previous exam

From image symmetry filters, which were based on weighted coefficients of the adapted Walsh-Hadamard transformation, we defined a summary measure for the symmetry of fibroglandular tissue for an individual image, *I*, as follows:

$$\mathrm{S}(I)=\sum_i w_i\left[\log S_i(I)\right]^2,$$

where S_i are tissue symmetry filters. This measure was evaluated by comparing values from contralateral CC pairs from individual women to pairs of randomly selected CC views.

3. RESULTS AND DISCUSSION

A histogram of the three distributions of measure of similarity in tissue pattern, is shown in Figure 6. It appears that, by our measure of similarity, contralateral views from individual women tend to be somewhat more similar than pairs consisting of a current and a prior view, though the informal nature of the study does not allow one to draw conclusions about the significance of this observation. The distribution of randomly selected pairs, which is shifted toward the dissimilar end of the scale, can be clearly distinguished from both of the previous distributions, though occasionally randomly selected images appear to be similar.

Figure 7 is a scatter plot with each point indicating the two symmetry values for a pair of images. The squares are for left and right CC views from the same woman and the circles are for pairs of CC views from different women. The contralateral pairs are clearly much more highly correlated than the randomly selected views, which appear to be uncorrelated.

These results are of a very preliminary nature and were obtained without the rigid controls that would be required for a more definitive study. This seems to be appropriate in this case because the methods being investigated are still under development. Nevertheless, these studies indicate that the approach we are taking has potential for detecting global differences between images.

We realize that there are more specific ways to measure differences in the image properties considered in this paper. However, measuring each property would require the application of specific techniques tailored to the property. The method considered here makes it possible to measure a wide range of properties within a very general framework.

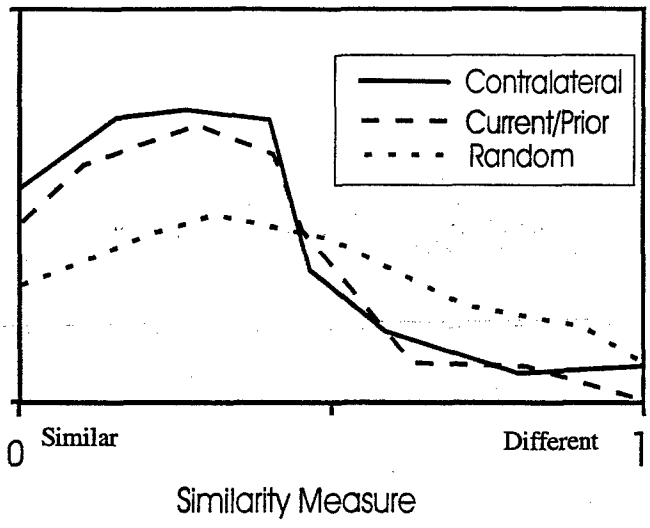


Figure 6 Histogram of similarity measure for three classes of image pairs

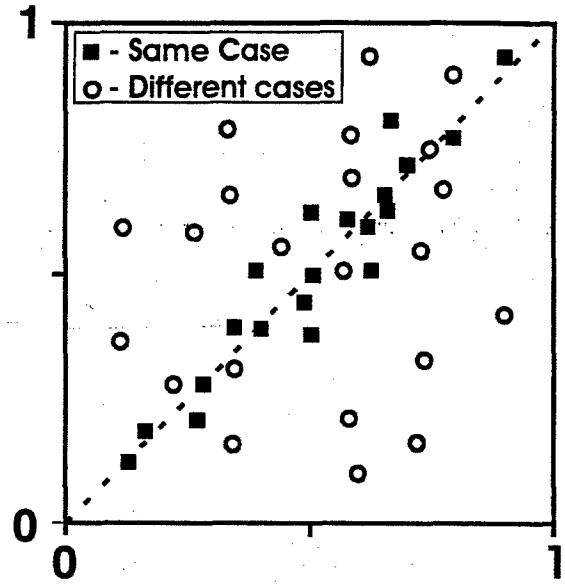


Figure 7 Scatter plot showing symmetry values for two classes of image pairs

4. CONCLUSION

This project proposes combining several techniques to produce an automated system for analyzing differences between mammographic images, and as such, it should help clarify the question of how much information can be derived from a comparison of contralateral images or a comparison of current and prior exams. The approach is unique in that it does not rely on the subtraction of mammographic views, or on the local geometric congruence of images.

ACKNOWLEDGEMENTS

This work is sponsored in part by grant IMG2000362 from the Susan G. Komen Foundation and grant CA80836 from the National Cancer Institute, National Institutes of Health, and also by the US Army Medical Research Acquisition Center, 820 Chandler Street, Fort Detrick, MD 21702-5014 under Contract DAMD17-02-1-0549. The content of the contained information does not necessarily reflect the position or the policy of the government, and no official endorsement should be inferred.

REFERENCES

1. B Zheng, YH Chang, D Gur. Computerized detection of masses from digitized mammograms: Comparison of single-image segmentation and bilateral image subtraction. Acad Radiol 1995; 2:1056-106

2. Difazio MC, MacMahon H, Xu XW, Tsai P, Shiraishi J, Armato SG, Doi K. Digital chest radiography: effect of temporal subtraction images on detection accuracy. *Radiology* 1997; 202:447-452.
3. FF Yin, ML Giger, K Doi, CE Metz, CJ Vyborny, RA Schmidt. Computerized detection of masses in digital mammograms: analysis of bilateral subtraction images. *Med Phys* 1991; 18:955-96
4. FF Yin, ML Giger, CJ Vyborny, K Doi, RA Schmidt. Comparison of bilateral-subtraction and single-image processing techniques in the computerized detection of mammographic masses. *Invest Radiol* 1993; 6:473-481.
5. FF Yin, ML Giger, K Doi, CJ Vyborny, RA Schmidt. Computerized detection of masses in digital mammograms: Automated alignment of breast images and its effect on bilateral-subtraction technique. *Med Phys* 1994; 21:445-452.
6. Mendez AJ, Tahocas PG, Loda MJ. Computer-aided diagnosis: automatic detection of malignant masses in digital mammograms. *Med Phys* 1998; 25:957-964.
7. M Sallam, K Bowyer. Registering time sequences of mammograms using a two-dimensional image unwarping technique. SIWDM 1994.
8. D Brzakovic, N Vujovic, M Neskovic, P Brzakovic, K Fogarty. Mammogram analysis by comparison with previous screenings. SIWDM 1994.
9. S. Sanjay-Gopal, HP Chan, N Petrick, TE Wilson, B Sahiner, MA Helvie, MM Goodsitt. A regional mammogram registration technique for automated analysis of interval changes of breast lesions. Proc. SPIE 3338:118-123.
10. Wang XH, Zheng B, Chang H, Good WF. Correction of digitized mammograms to enhance soft display amd tissue composition measurement. *Proc SPIE on Medical Imaging* 2001; Vol 4319.
11. U. Bick, M. L. Giger, R. A. Schmidt, R. M. Nishikawa, and K. Doi, "Density correction of peripheral breast tissue on digital mammograms," *RadioGraphics* 16, pp. 1403-1411, 1996.
12. J. W. Byng, J. P. Critten, and M. J. Yaffe, "Thickness-equalization processing for mammographic images," *Radiology* 203, pp. 564-568, 1997.
13. Stefanoyiannis AP, Costaridou L, Sakellaropoulos P. A digital density equalization technique to improve visualization of breast periphery in mammography. *British J Radiol* 2000; 73:410-420.
14. Good WF, Zheng B, Chang YH, Wang XH, Maitz G. Generalized procrustean image deformation for subtraction of mammograms. *Proc SPIE on Medical Imaging* 1999; 3666-167.
15. Good WF, Zheng B, Chang YH, Wang XH, Maitz G. Image modification for displaying temporal sequences of mammograms. *Proc SPIE on Medical Imaging* 2000; 3976 (in press).
16. Good WF, Zheng B, Chang YH, Wang XH, Maitz GS, Gur D. Multi-image CAD employing features derived from ipsilateral mammographic views. *Proc SPIE on Medical Imaging* 1999; 3661-47.
17. Hadamard J. Resolution d'une Question Relative aux Determinants. Bull.Sci.Math. Ser.2, 17, Part 1, 1893; 240-246.
18. Harmuth HF. Transmission of Information by Orthogonal Functions. Second Edition, Springer-Verlag, New York, 1972.